INTERTEMPORAL CHOICE AND CONSUMPTION MOBILITY

Tullio Jappelli
University of Salerno CSEF, and CEPR

Luigi Pistaferri
Stanford University, CEPR, and SIEPR

Abstract
The theory of intertemporal consumption choice makes sharp predictions about the evolution of the entire distribution of household consumption, not just about its conditional mean. In the paper, we study the empirical transition matrix of consumption using a panel drawn from the Bank of Italy Survey of Household Income and Wealth. We estimate the parameters that minimize the distance between the empirical and the theoretical transition matrix of the consumption distribution. The transition matrix generated by our estimates matches remarkably well the empirical matrix, both in the aggregate and in samples stratified by education. Our estimates strongly reject the consumption insurance model and suggest that households smooth income shocks to a lesser extent than implied by the permanent income hypothesis. (JEL: D52, D91, I30)

1. Introduction

How much consumption responds to income shocks is a central question in macroeconomics. As is well known, this response depends on the particular model that characterizes consumption behavior. For instance, the permanent income hypothesis (PIH) implies that households set consumption equal to permanent income, smoothing out transitory income fluctuations, but responding one-to-one to permanent shocks. On the other hand, under perfect insurance markets the distribution of marginal utilities is constant over time, and consumption does not respond to idiosyncratic income shocks. Deaton and Paxson (1994) have studied the implications of these theories for the dynamics of consumption inequality. They show that under the PIH or other models with incomplete markets, consumption inequality within a group of households with fixed membership should, on

Acknowledgments: Thanks are due to three anonymous referees, the Editor Roberto Perotti, and to Orazio Attanasio, Chris Carroll, Miles Kimball, Raffaele Miniaci, John Pencavel, and seminar participants at the 2004 Meeting of the Society of Economic Dynamics in Florence, Padua, 2000 World Congress of the Econometric Society in Seattle, and Michigan for comments. Work supported in part by the European Community's Human Potential Programme under contract HPRN-CT-2002-00235 (AGE: The Economics of Aging in Europe), the "Taube Faculty Research Fund" at the Stanford Institute for Economic and Policy Research, and the Italian Ministry of Education, Universities and Research (MIUR).

E-mail addresses: Jappelli: tullioj@tin.it; Pistaferri: pista@leland.standford.edu

Journal of the European Economic Association March 2006 4(1):75–115

average, increase with age. By contrast, the cross-sectional variance of marginal utilities should be constant over time under perfect insurance markets. The dynamics of consumption inequality is therefore informative about the impact of income shocks and about the validity of models of consumption behavior.[1]

Our point of departure from this literature is that measures of consumption inequality do not always provide an accurate measure of household behavior. Consumption inequality is a static concept, and as such it cannot handle violations of the life cycle–permanent income hypothesis (such as borrowing constraints or myopic behavior), which would imply a role for transitory income fluctuations over and above permanent fluctuations, or buffer stock behavior, which would imply more smoothing of income shocks than predicted by the standard model. The handling of these issues, we argue, calls for an analysis of consumption mobility.

The distinction between consumption inequality and consumption mobility is, effectively, a distinction between static and dynamic features of a distribution. Inequality refers to the dispersion of consumption at a point in time. Mobility describes movements within the consumption distribution as time goes by. Studies of consumption inequality may record no change in the dispersion of the underlying distribution even in the presence of intradistributional movements, with direct implications for welfare analysis.[2] Despite the importance of these issues, to the best of our knowledge the present paper represents the first attempt to analyze consumption mobility, both theoretically and empirically.[3] As we shall see, the analysis of consumption mobility delivers new implications of various theoretical models of intertemporal choice, generates new empirical tests and insights of those models, and allows estimation of the impact of income shocks on consumption.

The paper attempts to understand which model of intertemporal consumption choice is capable of explaining the amount of consumption mobility we observe in the data. We focus on several consumption theories, among which the theory of consumption insurance, the rule-of-thumb model, and the PIH model have

1. Cutler and Katz (1992) provide descriptive analysis of consumption inequality. More recently, motivated by the increase in income inequality in the U.S., a strand of papers has analyzed patterns of consumption inequality. Heathcote, Storesletten, and Violante (2003), for example, explore the implications of the recent sharp rise in U.S. wage inequality for welfare and the cross-sectional distributions of hours worked, consumption, and earnings.

2. In contrast to the analysis of consumption, there is a long tradition of studies of earnings and income mobility. Existing contributions can be divided into two broad groups. A first group analyzes transition probabilities across quantiles of the earnings distribution by Markov-chain models (e.g., Shorrocks 1978). A second approach is to specify and estimate a process for the conditional mean of earnings (e.g., Lillard and Willis 1978).

3. Phelan (1994) constructs transition matrices for consumption and leisure for the purpose of characterizing the properties of constrained-efficient allocations in dynamic economies with private information. Krueger and Perri (2003) estimate the persistence of consumption using the second largest eigenvalue of the transition matrix of consumption. They do not consider the implications of their theoretical models for consumption mobility.

received the widest attention. We nest these popular consumption models and estimate the parameters that minimize the distance between the empirical and the theoretical transition matrix of the consumption distribution. The exercise is performed constructing a transition matrix for consumption and testing different hypotheses concerning consumption dynamics. Because to measure mobility one needs to follow households over time, the empirical analysis is conducted on a panel drawn from the Bank of Italy Survey of Household Income and Wealth for the years 1987 to 1995. The survey we use is representative of the Italian population, contains a measure of total nondurable consumption, and has good-quality income data. Because there are virtually no panel data sets with broad consumption measures, a by-product of this paper is to bring the data set to the attention of empirical macroeconomists.

To see how the theory of intertemporal choice delivers implications for consumption mobility, consider first the extreme case of full consumption insurance. According to this theory, the cross-sectional distribution of the marginal utility of consumption of any group of households is constant over time. Of course aggregate consumption can increase or decrease, so the growth of the marginal utility for any household can be positive or negative, but the relative position of each household in the cross-sectional distribution of marginal utilities does not change over time. Consumption insurance therefore makes strong predictions about the entire distribution of marginal utilities, not just its mean or variance. In particular, consumption insurance implies absence of mobility of the marginal utility of consumption between any two time periods, regardless of the nature of the individual income shocks and the time frame considered. If one observes people moving up and down in the distribution of marginal utilities one must therefore conclude that some people are not insulated from idiosyncratic shocks, a contradiction of the consumption insurance hypothesis.[4]

A second case we consider is the rule-of-thumb model which predicts that households simply set consumption equal to income in each period. Given that any change in current income translates into an equivalent change in consumption, one should expect a relatively high degree of consumption mobility if shocks are not correlated with the rank position in the initial distribution of consumption.

In more realistic models with incomplete markets and insurance opportunities, individuals use saving as a self-insurance device and are able to smooth away at least some of the income variability. In these models the extent of consumption smoothing depends, among other things, on market imperfections arising from informational or enforceability problems and on the presence of liquidity constraints. Within this class of models, the best known is the PIH, in which income shifts over time because of transitory (e.g., mean reverting) and permanent

4. Although this implication of consumption insurance was mentioned in a theoretical paper by Banerjee and Newman (1991), to our knowledge it has never been explored empirically.

(e.g., persistent or nonmean reverting) shocks. If people behave according to the PIH, consumption reacts mostly to permanent unanticipated income shocks but is almost insensitive to transitory ones. Households will therefore move up and down in the consumption distribution only in response to permanent shocks. Thus, in the absence of taste shocks or measurement error, one should expect a degree of mobility that is intermediate between the level predicted by the consumption insurance hypothesis and the rule-of-thumb model.

The rest of the paper is organized as follows. Section 2 presents the data and the empirical transition matrix of the marginal utility, measured with log consumption per capita. The empirical matrix turns out to be quite robust to various adjustments and controls for the effect of family size and labor supply on preferences. In Section 3 we review the implications for consumption dynamics of the theories of intertemporal consumption choice and consider how to account for measurement error in consumption and taste shocks in the utility function. In Section 4 we estimate the parameters of the consumption rule by minimizing the distance between the empirical and the simulated transition matrix of the consumption distribution. The results, presented in Section 5, reject statistically each of the simple representations of the consumption decision rule, and reveal that households smooth income shocks to a lesser extent than implied by the PIH. The estimated parameters are also able to reproduce remarkably well the difference in consumption mobility that we observe in samples stratified according to education. The results are robust to the presence of measurement error in income and remain unchanged if we estimate simultaneously the parameters of the income process and of the consumption rule. Section 6 summarizes our results.

2. Measuring Consumption Mobility

The first step of our analysis is to construct an empirical transition matrix of consumption. This requires longitudinal household data. For this purpose we use the 1987–1995 panel of the Italian Survey of Household Income and Wealth (SHIW). This data set contains measures of consumption, income, and demographic characteristics of households. The SHIW provides a measure of total nondurable consumption, not just food, thus overcoming one of the main limitations of other panels, such as the PSID, that have been used to test for intertemporal consumption choice.

The SHIW, conducted by the Bank of Italy, surveys a representative sample of the Italian resident population. Sampling is in two stages, first municipalities and then households. Municipalities are divided into 51 strata defined by 17 regions and 3 classes of population size (more than 40,000; 20,000 to 40,000; less than 20,000). Households are randomly selected from registry office records.

From 1987 through 1995 the survey was conducted every other year and covered about 8,000 households, defined as groups of individuals related by blood, marriage, or adoption and sharing the same dwelling. Starting in 1989, each SHIW has re-interviewed some households from the previous surveys. The panel component has increased over time: 15% of the sample was re-interviewed in 1989, 27% in 1991, 43% in 1993, and 45% in 1995.[5] The response rate (ratio of responses to contacted households net of ineligible units) was 25% in 1989, 54% in 1991, 71% in 1993, and 78% in 1995.[6] Although these figures uncover considerable sample attrition especially in the early years of the survey, they are comparable to those obtained in other microeconomic data sets. For instance, in 1994 the net response rate in the U.S. Consumer Expenditure Survey was 83% for the interview sample and 81% for the diary sample. Given the rotating sample structure, the number of repeated observations on households in our sample ranges from a minimum of two to a maximum of five. Ample details on sampling, response rates, processing of results, and comparison of survey data with macroeconomic data are provided by Brandolini and Cannari (1994).

The total number of consumption transitions—that is, observations for the same household that are repeated over two adjacent years of data—is 10,508. To minimize measurement error we exclude cases in which the head changes over the sample period or gives inconsistent age figures. In most cases, the excluded households are those facing breaking-out events (widowhood, divorce, separation, etc.), leading to changes in household head. Inconsistent age figures can reflect unrecorded change in household head or measurement error. After these exclusions, the sample has 9,214 consumption and income transitions. Consumption is the sum of all expenditure categories except durables. Income is defined as the sum of labor income and transfers of all household members, excluding income from assets. These are the standard consumption and income concepts used in studies that test the implications of the permanent income hypothesis.[7]

Table 1 reports sample statistics of log consumption, income, and other household characteristics. All statistics are computed using sample weights. The panel is relatively stable over the sample period. Consumption grows considerably between 1987 and 1989 and is roughly constant afterwards. Over time,

5. In the panel component, the sampling procedure is also determined in two stages: (i) selection of municipalities (among those sampled in the previous survey); (ii) selection of households reinterviewed. This implies that there is a fixed component in the panel (for instance, households interviewed five times between 1987 to 1995, or four times from 1991 to 1995) and a new component every survey (for instance, households re-interviewed only in 1989).

6. Response rates increase in 1991 because in that year households included in the panel were chosen among those that had previously expressed their willingness to being re-interviewed.

7. Adding back asset income or asset income net of imputed rents does not change the main results of the paper.

TABLE 1. Descriptive statistics.

	1987	1989	1991	1993	1995	All years
ln c_t	9.90	10.08	10.02	10.01	10.00	10.02
var (ln c_t)	0.26	0.26	0.29	0.29	0.27	0.28
Gini coefficient of c_t	0.28	0.27	0.29	0.29	0.27	0.28
ln y_t	10.25	10.40	10.36	10.27	10.27	10.32
var (ln y_t)	0.39	0.37	0.37	0.57	0.47	0.45
Gini coefficient of y_t	0.35	0.32	0.32	0.36	0.35	0.34
South	0.41	0.37	0.34	0.36	0.39	0.37
North	0.43	0.46	0.48	0.47	0.43	0.46
Family size	3.15	3.12	3.04	3.07	3.01	3.07
Self-employed	0.20	0.17	0.17	0.16	0.15	0.16
Years of schooling	7.38	7.97	8.19	8.03	8.10	8.03
Less well-educated	0.78	0.72	0.70	0.72	0.70	0.72
More educated	0.22	0.28	0.30	0.28	0.30	0.28
Age	52.00	52.52	52.78	53.05	55.03	53.22
Born $\leq$ 1940	0.60	0.58	0.54	0.49	0.49	0.53
Born > 1940	0.40	0.42	0.46	0.51	0.51	0.47
Income recipients	1.63	1.72	1.72	1.74	1.78	1.73
Number of observations	1,097	2,717	4,036	4,006	3,211	15,067

Note: Cross-sectional means and variances are computed using sample weights. The variables c_t and y_t denote household nondurable consumption and disposable income, respectively. Demographic characteristics refer to the household head.

family size declines while the number of income recipients increases. Other demographic characteristics remain roughly unchanged. Self-employment falls slightly over time. Income strongly declines in 1993, a recession year, while dispersion increases. In all years, household disposable income is more variable than consumption. Note also the stability of the cross-sectional variance of log consumption as opposed to the wide fluctuations in the cross-sectional variance of log income. The pattern of the Gini coefficients for consumption and income confirms that the income distribution is less equal than the consumption distribution (34% vs. 28%). Interestingly, the 1993 recession boosts income inequality while leaving consumption inequality unaffected. These descriptive statistics are consistent with models in which households are able to smooth away at least some of the income shocks.

The focus of the present analysis, however, is not consumption inequality but consumption mobility. In what follows, we focus on the transition matrix of the logarithm of nondurable per capita consumption ($\log(c)$). We also check the sensitivity of the results using different consumption equivalent scales and the interactions between consumption and labor supply (see subsequent discussion). Note that the transition matrix estimated for $\log(c)$ is identical to that estimated for $c^{-\gamma}$, where γ is a constant. Thus, if individual utility is isoelastic ($u(c) = (1-\gamma)^{-1}c^{1-\gamma}$), and in the absence of taste shocks and measurement error in consumption, the empirical transition matrix of log consumption can also be interpreted as the empirical transition matrix of marginal utilities, a convenient feature in the light of our empirical strategy below.

2.1. The Empirical Transition Matrix

Assume that **P** is an unobservable $q \times q$ stochastic transition matrix of log household consumption, q being the number of consumption classes in the distribution. These classes could be determined exogenously or estimated from the quantiles of the empirical distribution. For notational simplicity we consider transition probabilities from period t to period t + 1; extending the argument to transition probabilities in periods $t + 2$, $t + 3$, and so on, is straightforward. The generic element of **P** is p_{ij}, the probability of moving from class i in period t to class j in period $t + 1$ conditioning on being in class i in period t. Define n_{ij} as the number of households that move from class i in period t to class j in period $t + 1$, $n_i = \sum_{i=1}^{q} n_{ij}$ as the total number of observations in each row i of **P**, and $n = \sum_{i=1}^{q} n_i$ the total number of observations. The maximum likelihood estimator of the first-order Markov transition probabilities is $\hat{p}_{ij} = n_{ij}/n_i$ (Anderson and Goodman 1957).[8]

Table 2 reports the transition matrix of log per capita consumption from 1987–1989 to 1993–1995. Recall that the SHIW is conducted every 2 years, so we observe transitions from period $t - 2$ to period t. The elements of the main diagonal report the proportion of households that did not change quartile. For instance, the entry in the top left cell of the 1993–1995 panel indicates that 68% of the households in the first quartile in 1993 were still in that quartile 2 years later. Off-diagonal elements signal consumption mobility. For instance, the second entry in the first row indicates that 25% of households moved from the first quartile in 1993 to the second quartile in 1995. The transition matrices for other years are similar, displaying substantial amount of consumption mobility.[9]

Each of the transition matrices is based on the distribution of log per capita consumption and does not take into account the fact that household expenditures might be affected by demographic variables and labor supply choice, breaking the link between the dynamics of log consumption and that of the marginal utility of consumption.

Changes in family size, for instance the arrival of children, alter family needs, hence, consumption allocations. Similarly, if household expenditures are

8. There are two methods for constructing the empirical counterpart of **P**. One is to keep the width of the consumption interval constant and let the number of observations within each interval vary. The alternative, more standard method, is to keep constant the marginal probabilities and let the interval width change, for instance dividing the distribution into discrete quantiles. We proceed using quartiles throughout; results with deciles are qualitatively similar and are not reported, for brevity.

9. In the simulation analysis we will make the assumption that consumption mobility is generated by a symmetric distribution of income shocks, which implies that our simulated transition matrix is also symmetric. It is therefore of interest to check if the transition matrix is symmetric using the maximum likelihood test suggested by Bishop, Fienberg, and Holland (1975). The statistic is of the form $\Psi = \sum_{i>j}(p_{ij} - p_{ji})^2/(p_{ij} + p_{ij}) \sim \chi^2_{q(q-1)/2}$. The p-value of the test is close to 1 for all years and does not reject the null hypothesis of symmetry.

TABLE 2. The transition matrix of consumption.

1987–1989				
	1989 Quartile			
1987 Quartile	1st	2nd	3rd	4th
1st	0.71	0.20	0.07	0.02
2nd	0.23	0.42	0.27	0.08
3rd	0.08	0.29	0.40	0.23
4th	0.03	0.09	0.29	0.60
1989–1991				
	1991 Quartile			
1989 Quartile	1st	2nd	3rd	4th
1st	0.66	0.25	0.07	0.01
2nd	0.25	0.41	0.27	0.06
3rd	0.10	0.27	0.41	0.25
4th	0.01	0.07	0.25	0.68
1991–1993				
	1993 Quartile			
1991 Quartile	1st	2nd	3rd	4th
1st	0.63	0.26	0.08	0.02
2nd	0.23	0.38	0.29	0.09
3rd	0.11	0.28	0.37	0.25
4th	0.04	0.10	0.26	0.60
1993–1995				
	1995 Quartile			
1993 Quartile	1st	2nd	3rd	4th
1st	0.67	0.25	0.07	0.01
2nd	0.24	0.43	0.27	0.07
3rd	0.07	0.27	0.43	0.23
4th	0.02	0.05	0.23	0.69

Note: The table reports consumption transitions from period $t-2$ to period t. The generic element of this table is $\widehat{p}_{ij}$, the estimated probability of moving from quartile i in period $t-2$ to quartile j in period t. Define n_{ij} as the number of households that move from quartile i in period $t-2$ to quartile j in period t and $n_i = \sum_j n_{ij}$ as the total number of observations in each row i of the transition matrix. The maximum likelihood estimator of the first-order Markov transition probabilities is then $\widehat{p}_{ij} = n_{ij}/n_i$.

characterized by economies of scale, one would observe mobility in consumption per capita even if the distribution of consumption per adult equivalent is constant over time. We thus compute transitions using log consumption per adult equivalent rather than per capita.[10] The pattern of the transition matrix and of the

10. There is a large literature on the cost of children and on the economies of scale in consumption, see Deaton (1997) for a survey. Any particular choice of an equivalence scale is therefore to a certain extent arbitrary, depending on the estimation method and assumptions about the utility function. We rely on a plausible equivalence scale that is consistent with current literature, assigning a weight of 1 to the first adult, 0.8 to any additional adult, and 0.25 to each household member less than 18 years old. We obtain similar results changing the parameters of the equivalence scale within a range of realistic estimates (0.1 to 0.5 for children, 0.6 to 1 for adults).

associated mobility index is unaffected. As a further check, we restrict attention to households whose demographic structure did not change over the sample period and find, again, similar consumption transitions.[11] In the remainder of the paper we thus focus on log consumption per capita.

If leisure is an argument of the utility function, and if consumption and leisure are nonseparable, consumption decisions are affected by predictable changes in households' labor supply (Attanasio 2000). This implies that the dynamics of consumers' rank in the consumption distribution depends, among other things, on changes in hours of work. Failure to control for changes in labor supply might therefore induce consumption mobility even in the absence of income and other idiosyncratic shocks. The interaction between consumption and labor supply is unlikely to affect the transition matrix of log consumption, however. First of all, in our sample, hours worked by individual household members and the proportion of spouses working do not change appreciably over the period considered. Second, even if we exclude households reporting changes in labor force participation (which induce the most dramatic shifts in labor supply), the consumption transition matrix is almost identical to the full sample matrix. For instance, when we apply this restriction, the main diagonal elements of the 1993–1995 transition matrix are 0.67, 0.42, 0.43, 0.69 (instead of 0.67, 0.43, 0.43, 0.69). These robustness checks suggest that in our sample the transition matrix of marginal utilities is not affected by demographic changes or by the labor supply status of the household.

2.2. Inference

The transition matrices in Table 2 are informative about the amount of consumption mobility—or the mobility of marginal utility—that we observe in the data. In the next section we derive from theory the implications for consumption mobility of popular models of intertemporal choice (PIH, rule-of-thumb, and consumption insurance), nest the models in a convenient unifying framework, and confront them with the data. Then, in Section 4 we discriminate among these models by estimating the parameters that minimize the distance between the empirical and the theoretical transition matrix of the consumption distribution.

To make inference about the empirical transition matrix, we will rely on the modified χ^2 goodness-of-fit statistic proposed by Anderson and Goodman (1957):

$$\sum_{i=1}^{q}\sum_{j=1}^{q} n_i \frac{(\hat{p}_{ij} - p_{ij}^0)^2}{\hat{p}_{ij}} \sim \chi^2_{q(q-1)}. \tag{1}$$

11. For instance, excluding households with changes in family composition results in a mobility index of 0.576 in 1993–1995.

The statistic compares estimated and theoretical transition probabilities and can be used to test the null hypothesis that $p_{ij} = p_{ij}^0$ for all i, j. A similar statistic can be used to test if the transition matrix differs statistically over time or between population groups:

$$\sum_{i=1}^{q}\sum_{j=1}^{q} N_i \frac{\left(\hat{p}_{ij}^{g_0} - \hat{p}_{ij}^{g_1}\right)^2}{\hat{p}_{ij}} \sim \chi^2_{q(q-1)},$$

where $\hat{p}_{ij}$ is the estimate of p_{ij} obtained pooling data for two groups or time periods g_0 and g_1, and $N_i^{-1} = (1/n_i^{g_0}) + (1/n_i^{g_1})$.

In the empirical application we will also be interested in matching the empirical transition matrix with a simulated matrix that depends on a vector of unknown parameters θ. This estimation problem can be addressed by implementing a minimum χ^2 method, namely, minimizing the function

$$\sum_{i=1}^{q}\sum_{j=1}^{q} n_i \frac{\left(\hat{p}_{ij} - p_{ij}(\theta)\right)^2}{p_{ij}(\theta)}. \tag{2}$$

The properties of this estimator are discussed in Neyman (1949). In Appendix B we show that (2) can be rewritten as

$$(\hat{\mathbf{p}} - \mathbf{p}(\theta))\, \Omega\, (\theta)^{-1} (\hat{\mathbf{p}} - \mathbf{p}(\theta))', \tag{3}$$

where—after deleting a column from the theoretical and empirical transition matrices to avoid singularity—$\hat{\mathbf{p}}$ is the vector of estimated transition probabilities, $\mathbf{p}(\theta)$ the vector of theoretical transition probabilities, and $\Omega(\theta)$ the covariance matrix of the distance vector $(\hat{\mathbf{p}} - \mathbf{p}(\theta))$. The function (3) has therefore the optimal minimum distance form of Chamberlain (1984) that econometricians are familiar with.[12]

Neyman (1949) also proposed a modified minimum χ^2 method, where the function to minimize is

$$\sum_{i=1}^{q}\sum_{j=1}^{q} n_i \frac{(\hat{p}_{ij} - p_{ij}(\theta))^2}{\hat{p}_{ij}}.$$

Appendix B proves that this function can be rewritten as $(\hat{\mathbf{p}} - \mathbf{p}(\theta))\, \hat{\Omega}^{-1} (\hat{\mathbf{p}} - \mathbf{p}(\theta))'$, where $\hat{\Omega}$ uses the estimated $\hat{p}_{ij}$ to construct an estimate of the covariance matrix of $(\hat{\mathbf{p}} - \mathbf{p}(\theta))$.

12. Note that when $\dim(\hat{\mathbf{p}}) > \dim(\theta)$, the minimized value of (3) can also be interpreted as the statistic for testing overidentifying restrictions. The statistic is distributed $\chi^2_{\dim(\hat{\mathbf{p}})-\dim(\theta)}$.

When the expression for $p_{ij}(\theta)$ is available in closed form, implementation of the minimum χ^2 criterion is straightforward. When it is not, as in our case, one must rely on simulations to generate the transition probability conditional on θ, and then apply the minimum χ^2 method to the simulated $p_{ij}(\theta)$. Details are provided in Section 4 and in Appendix A.

3. Intertemporal Choice and Mobility

To explore the relation between the consumption and the income distributions, it is useful to start by presenting a fairly general characterization of the income process. Starting with Hall and Mishkin (1982), it has become quite standard in panel data studies of income and consumption dynamics to express log income of household h in period t as

$$\ln y_{h,t} = \beta X_{h,t} + p_{h,t} + e_{h,t}, \tag{4}$$

where $X_{h,t}$ is a set of deterministic variables such as age and region of residence, and $p_{h,t}$ and $e_{h,t}$ are permanent and transitory components, respectively.[13] The latter is the sum of an idiosyncratic ($\varepsilon_{h,t}$) and an aggregate component (ε_t); both are assumed to be serially uncorrelated. Because the permanent component of income changes very slowly, the standard assumption is to model it as a random walk process of the form

$$p_{h,t} = p_{h,t-1} + z_{h,t}, \tag{5}$$

where $z_{h,t}$ is the permanent innovation, which is again the sum of an idiosyncratic ($\zeta_{h,t}$) and an aggregate shock (ζ_t); both components are serially uncorrelated. We also assume that $\varepsilon_{h,t}$ and $\zeta_{h,t}$ are mutually uncorrelated disturbances with variances σ_ε^2 and σ_ζ^2, respectively. Because we operate with a short panel, transitory and permanent aggregate shocks will be estimated by a vector of time dummies, d_t.

The decomposition of income shocks into transitory and permanent components dates back to Friedman (1957). Some of the income shocks are transitory (mean reverting) and their effect does not last long. Examples include fluctuations in overtime labor supply, bonuses, lottery prizes, and bequests. On the other hand, some of the innovations to earnings are highly persistent (nonmean reverting) and their effect cumulates over time. Examples of permanent innovations are generally associated with job mobility, promotions, layoffs, and severe health shocks.

Given our assumptions, income growth can be written as

$$\Delta \ln y_{h,t} = \Delta d_t + \beta \Delta X_{h,t} + \zeta_{h,t} + \Delta\varepsilon_{h,t}. \tag{6}$$

13. The logarithmic transformation eliminates heteroskedasticity in the distribution of income in levels.

As we shall see, this income process delivers different implications for consumption mobility for different models of intertemporal choice. We also consider how these implications change in the presence of taste shocks, measurement error in consumption, and measurement error in income.

3.1. The Permanent Income Hypothesis

The first model we consider is a version of the PIH with CRRA preferences $u(c) = (1-\gamma)^{-1}c^{1-\gamma}$, where γ^{-1} is the elasticity of intertemporal substitution. We assume that infinitely lived households maximize expected utility under perfect credit markets, subject to an intertemporal budget constraint. We assume that income follows the process (4)–(5) and that it is the only source of uncertainty of the model. As in Blundell and Preston (1998) and Blundell, Pistaferri, and Preston (2003), we approximate the Euler equation for consumption with a second-order Taylor expansion and assume that $r = \delta$, that consumption equals permanent income, and that the conditional variance of income shocks varies only in the aggregate. One can show that under such assumptions, individual consumption growth depends on an aggregate component m_t^{PIH} and unanticipated idiosyncratic income shocks:

$$\Delta \ln c_{h,t} = m_t^{\mathrm{PIH}} + \frac{r}{1+r}\varepsilon_{h,t} + \zeta_{h,t}. \tag{7}$$

Equation (7) indicates that the optimal rule is to respond one-to-one to permanent shocks ζ and to revise consumption only by the annuity value of the income innovation in case of transitory shocks ε. This is in fact the basic insight of the PIH, where people self-insure against high-frequency income shocks but adjust their consumption fully in response to low-frequency shocks. As we shall see, a convenient feature of equation (7) is that it can easily be nested with consumption rules derived from different models.

Suppose now that we observe a given cross-sectional distribution of consumption at time $t-1$ and that the income shocks are not perfectly correlated with the consumption rank of each household in the cross-section. Because aggregate shocks are by definition identical for all households, they do not change each consumer's rank in the consumption distribution and therefore they will not induce any consumption mobility: If they were the sole source of consumption fluctuations the mobility index would be zero.[14] However, other shocks are idiosyncratic, and will move people up and down in the consumption distribution, to an extent that depends on the variance of the two shocks. But because the

14. Suppose that income shocks were instead perfectly and positively correlated with the rank of household consumption in the cross-section. Then, the poorest households receive the largest negative shocks and the richest the largest positive shocks, implying no mobility as in the consumption insurance case.

impact of transitory shocks is scaled down by the factor $r/(1+r)$, we expect the variance of the permanent shocks to have the greatest impact on mobility.

Simulation results produced by Carroll (2001) show that with constant relative risk aversion, impatient consumers, and an income process similar to the one we use, the implication of the PIH that transitory income shocks have a negligible impact on consumption still holds true. Permanent shocks, however, have a somewhat lower impact in buffer stock models. In fact, in such models permanent income shocks reduce the ratio of wealth to permanent income, thus increasing also precautionary saving. Under a wide range of parameter values, Carroll shows that in this class of models the marginal propensity to consume of a permanent income shock is about 0.9, not far from that of the approximation in (7). Therefore, empirically it is difficult to distinguish the PIH from buffer stock models on the basis of the marginal propensity to consume out of permanent income shocks. But the main intuition is still valid: If individuals smooth consumption and understand the income generating process, transitory income shocks should have a negligible impact on consumption. To account for the effect remarked by Carroll and others, in the empirical analysis we take into account the degree of consumption smoothing arising from precautionary savings estimating:

$$\Delta \ln c_{h,t} = m_t^{\mathrm{PIH}} + \phi \left(\frac{r}{1+r} \varepsilon_{h,t} + \zeta_{h,t} \right).$$

In buffer stock models ($\phi < 1$), assets accumulated for precautionary purposes allow people to smooth income shocks to a larger extent than in the PIH model ($\phi = 1$).

3.2. The Rule-of-Thumb Model

Let us assume that consumption equals income in each period, that is,

$$\ln c_{h,t} = \ln y_{h,t}.$$

This model has been often proposed as a simple, yet extreme alternative to the PIH to describe the behavior of households that do not use savings to buffer income shocks but spend all they receive. Some authors rationalize this model by appealing to the presence of binding liquidity constraints in each period. Laibson (1997) shows that it is the equilibrium outcome for hyperbolic (or *golden-eggs*) consumers.[15] We term it rule-of-thumb model to indicate a situation in

15. In the hyperbolic consumer model, individuals have preferences that change over time (there are different selves in different periods). In the model proposed by Laibson (1997) *self* $t-1$ chooses assets a_{t-1} to constrain the consumption of *self* t. This is done by keeping most assets invested in an illiquid instrument. Hence, at any point in time, the consumer is effectively liquidity constrained, even though the constraint is self-imposed. Laibson shows that in equilibrium consumption is exactly equal to the current level of cash flow, or total income.

which consumption tracks income closely, even when individuals have accumulated assets in previous periods. The model is an interesting case to study because it approximates the behavior of consumers with short horizons, limited resources, or hyperbolic discount factors, giving an upper bound for the sensitivity of consumption to income shocks.

Using the income process (4)–(5), the dynamic of consumption is given by

$$\Delta \ln c_{h,t} = m_t^K + \varepsilon_{h,t} - \varepsilon_{h,t-1} + \zeta_{h,t}, \tag{8}$$

where m_t^K is the effect of the aggregate shocks on consumption in the rule-of-thumb model. According to the model, the growth rate of consumption is therefore equally affected by current and lagged transitory shocks and by permanent shocks. The main difference with the PIH is that in the rule-of-thumb model, transitory shocks impact one-to-one on consumption. It is precisely for this reason that in the rule-of-thumb model one should expect more consumption mobility than under the permanent income rule: There is another channel through which households move to a different quartile from one period to the next.

3.3. Consumption Insurance

To illustrate the implications of the theory of intertemporal choice with complete insurance markets, let us keep the assumption that households have preferences of the constant relative risk aversion (CRRA) type. The implications of the model are identical for any power utility function. As shown, among others, by Mas-Colell, Whinston, and Greene (1995), the optimal transition law for consumption with complete markets can be obtained by assuming that there is a social planner who maximizes a weighted sum of individual households' utilities. The Lagrangian of this problem can be written as

$$L = \sum_h \lambda_h \sum_s \sum_t \pi_{s,t} u\left(c_{h,s,t}\right) + \sum_s \sum_t \mu_{s,t} \left(C_{s,t} - \sum_h c_{h,s,t}\right),$$

where *h, s,* and *t* are subscripts for household *h* in the state of nature *s* in period *t*, λ_h is the social weight for household h, $\mu_{s,t}$ is the Lagrange multiplier associated with the resource constraint, $\pi_{s,t}$ the probability of the realization of state *s* in period *t*, and $C_{s,t}$ aggregate consumption in state *s* and period *t*.

The first-order condition can be written in logarithms as

$$-\gamma \ln c_{h,s,t} = \ln \mu_{s,t} - \ln \lambda_h - \ln \pi_{s,t}.$$

To obtain the growth rate of consumption, subtract side-by-side from the same expression at time $t-1$:

$$\Delta \ln c_{h,t} = -\gamma^{-1} \Delta \ln \mu_t + \gamma^{-1} \Delta \ln \pi_t \equiv m_t^{CI}, \tag{9}$$

where we drop the subscript s because only one state is realized in each period. The two terms on the right-hand side of equation (9) represent genuine aggregate effects. The first term is the growth rate of the Lagrange multiplier, the second is the growth rate of the state probabilities. Note that first-differencing has eliminated all household fixed effects (μ and π in equation 9 are not indexed by h).

Equation (9) states that the growth rate of the marginal utility of consumption (or of consumption, for that matter) of each household is the same. This implies that the initial cross-sectional distribution of consumption is a sufficient statistic to describe all future distributions. Because all households experience the same growth rate of marginal utility, and given our assumption of isoelastic utility, their rank in the consumption distribution is stationary. Note that the stationarity of the cross-sectional distribution is directly implied by the assumption that insurance markets fully insulate households from idiosyncratic shocks. The statistical counterpart of consumption insurance is that the transition matrix for household consumption is an identity matrix. The extreme assumptions of this model are clearly unrealistic. However, the model provides a lower bound for the impact of income shocks on consumption and is therefore a useful theoretical benchmark.

The discussion in Sections 3.1–3.3 suggests that under CRRA preferences (and no measurement error or taste shocks), consumption mobility should be zero in the consumption insurance model, intermediate in the permanent income model, and highest in the rule-of-thumb model. In practice, the presence of measurement error and/or taste shocks does not provide a clear-cut ordering of the models in terms of expected mobility. Nevertheless, it is still possible to distinguish between the different models by estimating the parameters that determine the extent of consumption mobility that we observe in the data.

3.4. Nesting the Three Models

The distinction between the three models is useful but too stylized for empirical applications. Consumption insurance is no less unrealistic than assuming that all income is consumed in each period, or that all households follow exactly the PIH. In the empirical application we therefore nest the three models and estimate the parameters of the following flexible consumption rule:

$$\ln c_{h,t} = \ln c_{h,t-1} + \phi \left(\frac{\lambda + r}{1 + r} \varepsilon_{h,t} - \lambda \varepsilon_{h,t-1} + \zeta_{h,t} \right). \tag{10}$$

Because aggregate shocks do not affect consumption mobility under each of the models discussed above, for notational simplicity equation (10) omits the aggregate component m_t^j. However, when we estimate the income process we control for aggregate shocks by introducing time dummies in the regression (d_t in equation (6)).

The two parameters λ and ϕ allow us to distinguish various forms of departure from the stylized models of intertemporal choice. Consider first the case in which $\phi = 1$. The parameter λ represents the extent to which consumption responds to income over and above the amount warranted by the PIH—that is, the excess sensitivity of consumption to current and past income shocks. One way to interpret this parameter is that each household sets consumption equal to income with probability λ (perhaps because of binding liquidity constraints or hyperbolic discounting) and follows the PIH with probability $(1 - \lambda)$. Note that with $\lambda = 0$ the expression (15) reduces to the PIH, while with $\lambda = 1$ one obtains the rule-of-thumb model where consumption equals income each period.

Consider now the situation in which $\phi = 0$. Income shocks play no role in the consumption insurance model. But intermediate cases in which $0 < \phi < 1$ are interesting and potentially informative, as discussed in Section 3.1. Some consumers have assets accumulated for precautionary reasons which allow them to smooth income shocks to a larger extent than in the PIH model (where $\phi = 1$).

3.5. Measurement Error and Taste Shocks

The consumption transition law is derived assuming that there is no measurement error in consumption. In practice, consumption mobility, as estimated from the data, is potentially upward-biased because of reporting errors. If respondents report their consumption with errors, one will find units moving up and down even if their true rank in the consumption distribution is unchanged; hence, measurement errors affect consumption dynamics and the extent of mobility we measure in the data. In the estimation it is therefore important to account explicitly for measurement error. We will do so by assuming that true consumption is measured with a multiplicative error:

$$\ln c^*_{h,t} = \ln c_{h,t} + \tilde{\omega}_{h,t}, \tag{11}$$

where $\ln c^*$ is measured consumption and $\tilde{\omega}$ is an independently and identically distributed measurement error.

As explained in Section 2, another source of consumption mobility is taste shocks. Even though in Section 2 the actual empirical matrix is quite stable across a wide range of robustness checks (such as defining consumption in terms of adult equivalents, or restricing the sample to minimize the potential impact of leisure on consumption transitions), the impact of unobserved taste shocks cannot be ruled out. To see what impact they have on the empirical estimates, we rewrite the utility function as $u(c, \xi) = (1 - \gamma)^{-1} c^{1-\gamma} \exp(\theta\tilde{\xi})$, where $\tilde{\xi}$ is an i.i.d. shock to individual preferences.

The combined presence of measurement error and taste shocks leads to the following reformulation of the consumption dynamics in equation (10):

$$\ln c^*_{h,t} = \ln c^*_{h,t-1} + \phi\left(\frac{\lambda + r}{1 + r}\varepsilon_{h,t} - \lambda\varepsilon_{h,t-1} + \zeta_{h,t}\right) + \left(\tilde{\omega}_{h,t} - \tilde{\omega}_{h,t-1}\right) + \psi\left(\tilde{\xi}_{h,t} - \tilde{\xi}_{h,t-1}\right), \tag{12}$$

where ψ is a function of θ, γ, and λ.

Without loss of generality, assume that in each period the standard deviation of measurement error ($\sigma_{\tilde{\omega}}$) is a fraction $\eta_{\tilde{\omega}}$ of the standard deviation of measured consumption, $\sigma_{\tilde{\omega}} = \eta_{\tilde{\omega}}\sigma_{\ln c*}$ and that the standard deviation of taste shocks ($\sigma_{\tilde{\xi}}$) is a fraction $\eta_{\tilde{\xi}}$ of the standard deviation of measured consumption, $\sigma_{\tilde{\xi}} = \eta_{\tilde{\xi}}\sigma_{\ln c*}$. Because the variables are expressed in logs, $\eta_{\tilde{\omega}}$ and $\eta_{\tilde{\xi}}$ can be interpreted as the percentage variability in observed consumption due to reporting error and taste shocks, respectively. Rewrite (12) as

$$\ln c^*_{h,t} = \ln c^*_{h,t-1} + \phi\left(\frac{\lambda + r}{1 + r}\varepsilon_{h,t} - \lambda\varepsilon_{h,t-1} + \zeta_{h,t}\right) + \eta_{\tilde{\omega}}\left(\omega_{h,t} - \omega_{h,t-1}\right) + \psi\eta_{\tilde{\xi}}\left(\xi_{h,t} - \xi_{h,t-1}\right), \tag{13}$$

where we have adopted the linear transformations $\omega_{h,\tau} = \eta_{\tilde{\omega}}^{-1}\tilde{\omega}_{h,\tau}$ and $\xi_{h,\tau} = \eta_{\tilde{\xi}}^{-1}\tilde{\xi}_{h,\tau}$, so that $\omega \sim$ i.i.d.$(0, \sigma_{\ln c*})$ and $\xi \sim$ i.i.d.$(0, \sigma_{\ln c*})$. It is clear from (13) that, in the absence of additional information, the separate effects of taste shocks and measurement error are not identified. We thus can define a composite error term $\tilde{\upsilon}_{h,t} = \eta_{\tilde{\omega}}\omega_{h,t} + \psi\eta_{\tilde{\xi}}\xi_{h,t}$, and rewrite (13) as

$$\ln c^*_{h,t} = \ln c^*_{h,t-1} + \phi\left(\frac{\lambda + r}{1 + r}\varepsilon_{h,t} - \lambda\varepsilon_{h,t-1} + \zeta_{h,t}\right) + \alpha\left(\upsilon_{h,t} - \upsilon_{h,t-1}\right), \tag{14}$$

where we have adopted the linear transformation $\upsilon_{h,\tau} = \alpha^{-1}\tilde{\upsilon}_{h,\tau}$, so that $\upsilon \sim$ i.i.d.$(0, \sigma_{\ln c*})$ and $\alpha = \left(\eta^2_{\tilde{\omega}} + (\psi\eta_{\tilde{\xi}})^2\right)^{1/2}$ is an unknown parameter to estimate. Equation (14) shows that the composite error term induces a further reason for consumption to vary. Clearly, not only consumption dynamics changes, but the implied consumption mobility as well.[16] In what follows, we will therefore interpret the parameter α as the amount of consumption variability induced by measurement error and taste shocks.

16. The clearest case in which this happens is in the model with consumption insurance (where $\phi = 0$) and no taste shocks: Measurement error can induce consumption mobility even though the growth rate of marginal utilities is constant.

4. Estimation Method

We now discuss estimation of the parameters of interest. One complication with the panel we use is that although income and consumption refer to calendar years, data are collected every other year from 1987 to 1995. The simulated transition laws for consumption must therefore be slightly modified to tackle this problem. One can use the recursive aspect of (14) to rewrite it as

$$\ln c^*_{h,t} = \ln c^*_{h,t-2} + \phi\left(\frac{\lambda + r}{1+r}\varepsilon_{h,t} + \frac{(1-\lambda)r}{1+r}\varepsilon_{h,t-1} - \lambda\varepsilon_{h,t-2} + \zeta_{h,t} + \zeta_{h,t-1}\right) + \alpha\big(\upsilon_{h,t} - \upsilon_{h,t-2}\big). \tag{15}$$

The parameters to be estimated are the variances of the permanent and transitory income shocks, the fraction of measurement error and taste shocks, the degree of excess sensitivity, the degree of income smoothing, and the real interest rate. As for the interest rate, we assume a value of 2% throughout.

The estimation of the income and consumption process can be performed in two steps. In the first step, we estimate the income variances σ^2_ε and σ^2_ζ. In the second step we use the estimated income variances to generate the income shocks ε and ζ and the composite error υ that appear in the consumption rule and estimate the parameters ϕ, λ, and α by simulated minimum χ^2 method.

As explained in Section 4, we specify the income process as

$$\ln y_{h,t} = d_t + \beta X_{h,t} + p_{h,t} + \varepsilon_{h,t},$$

where $y_{h,t}$ is per capita household disposable income and d_t a set of time dummies. Using the 1987–1995 panel, we regress ln $y_{h,t}$ on a set of demographic variables (north, south, a dummy for gender, a fourth-order age polynomial, and education dummies) and time dummies, so as to remove the deterministic component of income. We save the residuals $u_{h,t} = p_{h,t} + \varepsilon_{h,t}$ and carefully examine their covariance properties. We estimate covariances using equally weighted minimum distance methods, as suggested by Altonji and Segal (1997).[17]

We find that the estimated covariances are consistent with the income process in equations (4) and (5), namely, that there is a random-walk permanent component and a serially uncorrelated transitory shock. Recall that because of the sample design of the SHIW we can only construct the covariance matrix for two years apart income residuals, $u_{h,t} - u_{h,t-2} = \zeta_{h,t} + \zeta_{h,t-1} + \varepsilon_{h,t} - \varepsilon_{h,t-2}$. To check the consistency of the estimated income process with the model in equations (4)

17. Covariances can be estimated by equally weighted minimum distance or optimal minimum distance. As shown by Altonji and Segal (1996), the latter can produce inconsistent estimates in small samples, so we adopt the former.

and (5), note that the income process implies the following testable restrictions on the covariance matrix of the first difference of the income residuals:

$$E\left[\left(u_{h,\tau}-u_{h,\tau-2}\right)^2\right]=2\sigma_\zeta^2+2\sigma_\varepsilon^2$$
$$E\left[\left(u_{h,\tau}-u_{h,\tau-2}\right)\left(u_{h,\tau-2}-u_{h,\tau-4}\right)\right]=-\sigma_\varepsilon^2$$
$$E\left[\left(u_{h,\tau}-u_{h,\tau-2}\right)\left(u_{h,\tau-j}-u_{h,\tau-j-2}\right)\right]=0 \text{ for all } j\geq 4.$$

Provided that the restrictions are met in the data, one can estimate the variance of the transitory shock σ_ε^2 from the first-order autocovariance of income residuals and the variance of the permanent shock σ_ζ^2 combining information on the variance and the first-order autocovariance of the residuals. We find that the estimated autocovariance at the second order is very small (-0.0056) and not statistically different from zero (a t-statistic of -1.1); the autocovariance at the third order is again small (-0.0178) and not statistically different from zero (a t-statistic of -1.1). In contrast, the first-order autocovariance (which provides an estimate of $-\sigma_\varepsilon^2$) is precisely estimated (a t-statistic of 6.4) at -0.0794. The estimate of the overall variance ($2\sigma_\zeta^2+2\sigma_\varepsilon^2$) is 0.2122 (with a t-statistic of 19.4), so we infer that $\sigma_\zeta^2=0.0267$ and $\sigma_\varepsilon^2=0.0794$ (with standard errors 0.0135 and 0.0123, respectively).[18] These parameter estimates are broadly consistent with the evidence available for the U.S. where researchers have found variances of similar magnitude.[19]

The remaining unknown parameters are ϕ, the degree to which consumers are unable to insure income shocks through precautionary savings, λ, the degree of excess sensitivity of consumption, and α, the composite amount of consumption variability induced by measurement error and taste shocks. We therefore estimate ϕ, λ, and α minimizing the distance between the empirical and the theoretical transition matrix using the modified χ^2 criterion presented in Section 2.

Because theoretical transition probabilities do not have a closed form expression, we use a simulated minimum χ^2 estimation method.[20] A sketch of the estimation method is the following. We start by generating, for each household, draws for the transitory and the permanent income shocks and for the measurement error in consumption.[21] The income shocks are drawn from a normal distribution

18. Unfortunately, with data collected every 2 years we cannot distinguish between this income process and one where the transitory component is an MA(1) process.

19. For instance, Carroll and Samwick (1997), using the PSID, estimate $\sigma_\zeta^2=0.0217$ and $\sigma_\varepsilon^2=0.0440$.

20. Alan and Browning (2003) estimate the parameters of the Euler equation (the elasticity of intertemporal substitution and the amount of measurement error) using simulated Euler equation residuals. Gourinchas and Parker (2002) estimate the coefficient of relative risk aversion and the intertemporal discount rate using a method of simulated moments conditional on the assumption that the PIH is the true consumption model.

21. In each year we choose a sample size identical to the number of actual sample transitions (for instance, 2,982 in 1991–1993 and 3,211 in 1993–1995).

with mean zero and variances equal to the estimated variance from the income process ($\sigma_{\varepsilon}^2 = 0.0794$ and $\sigma_{\zeta}^2 = 0.0267$, respectively). The errors $\upsilon_{h,t}$ and $\upsilon_{h,t-2}$ are drawn from a normal distribution with mean zero and variance equal to the variance of measured log consumption at t and $t-2$, respectively. The number of draws is $S = 100$ for each household, for a total of HS simulated observations (H being the number of households). We then choose a starting value for the parameter vector and, for each household, compute next period consumption, $\ln c_{h,t}^*$, using (15). In this way, the covariance structure between individual income and individual consumption bears directly on the extent of consumption mobility. Finally, we compute the theoretical transition probabilities (averaging across the S simulations) and obtain the parameter estimates as those that minimize the (optimal) distance between empirical and theoretical transition probabilities. Because the number of transition probabilities that we fit exceeds the number of parameters we estimate, we have overidentifying information that can be used to assess the goodness of fit of the model. Appendix A reports technical details about the properties of this estimator and the minimization algorithm.

The two-step procedure described in this Section provides consistent, but not fully efficient, parameter estimates. In Section 5.4 we therefore check the sensitivity of the results estimating simultaneously income and consumption mobility, as well as the rank correlation statistics between the two variables.

5. Estimation Results

In this section we report full sample estimates of the parameters of the consumption rule and of the transition matrix for consumption. We also check the sensitivity of the estimates if we estimate simultaneously the parameters of the income process and of the consumption rule and when we consider measurement error in income. Finally, we split the sample by educational attainment of the head, estimate a separate income process for each education group, and evaluate patterns of consumption mobility of households with different levels of educational attainment.

5.1. Full Sample Estimates

The results of the full sample estimates are similar across periods, so we focus on the most recent one (1993–1995), which also features the largest number of transitions. The stability of the results across different sample periods suggests that the simulations are only marginally affected by the initial distribution of consumption (the income process and the associated variances of the shocks are in fact assumed to be the same across the different samples).

As a preliminary analysis, we constrain the parameter space in the simulated minimum χ^2 estimation method and compare the empirical and theoretical transition matrices in three benchmark models: PIH ($\phi = 1, \lambda = 0$), rule-of-thumb ($\phi = 1, \lambda = 1$), and consumption insurance ($\phi = 0$). In these experiments, we also rule out the effect of measurement error and taste shocks by setting $\alpha = 0$. These benchmark models illustrate our estimation strategy and provide a gateway to the results that follow.

We find that simulated consumption mobility, as measured by the Shorrocks mobility index, is highest in the rule-of-thumb model (65%), intermediate in the case of the permanent income hypothesis (44%), and zero under consumption insurance.[22] The ranking of mobility agrees with the discussion in Section 3.3 because idiosyncratic income shocks translate into consumption changes entirely in the rule-of-thumb model, partially in the PIH via intertemporal smoothing, and are fully insured in the risk sharing model. However, from a statistical point of view, none of these models is able to match the amount of empirical mobility. Cell-by-cell comparison of the theoretical and empirical transition matrices reveals that each of the three models is rejected according to the χ^2 goodness-of-fit statistics.[23]

To bridge the gap between simulated and empirical mobility we therefore consider the effect of measurement error and taste shocks in consumption, and allow for a more flexible response of consumption to income shocks than predicted by either full insurance, rule-of-thumb model, or PIH. We know from equation (15), nesting the three baseline models, that raising the excess sensitivity parameter λ or the insurance parameter ϕ also increases consumption mobility, regardless of the size of the composite error term.

We therefore implement the simulated minimum χ^2 estimation method freeing the parameter space. The parameter estimates of ϕ, λ, and α are reported in column (1) of Table 3. For completeness, we also report the estimates of σ_ζ^2 and σ_ε^2 obtained in the first step. Because the restriction $\phi = 1$ is not rejected statistically, we impose the restriction in column (2). The results indicate that the composite variability of consumption due to measurement error and taste shocks is 38% and that the excess sensitivity coefficient is 16%. Both estimates are precisely estimated and statistically different from zero at the 1% level. For values of $\phi = 1, \lambda = 0.16$ and $\alpha = 0.38$, the simulated mobility index is almost identical to the empirical one (60.13% vs. 59.37%). The χ^2 goodness-of-fit statistic (or test of overidentifying restrictions) is 15 with a p-value of 9%, indicating that the

22. The Shorrocks index, defined as $S(\mathrm{P}) = (q-1)^{-1}(q - \mathrm{trace}(\mathrm{P}))$, is a standard way of summarizing the extent of mobility from a transition matrix. See Shorrocks (1978).

23. In the rule-of-thumb case ($\alpha = 0, \lambda = \phi = 1$), the χ^2 value is 58, in the PIH ($\alpha = \lambda = 0$, $\phi = 1$) 250, and in the consumption insurance case ($\alpha = \lambda = \phi = 0$) 2,856. Each of these values exceeds the critical value of $\chi^2_{12;0.05} = 21$.

TABLE 3. Parameter estimates and χ^2 statistics.

	(1)	(2)
ϕ	0.9875	1.0000
	(0.0230)	
λ	0.1586	0.1622
	(0.0377)	(0.0248)
α	0.3875	0.3822
	(0.0186)	(0.0116)
Variance of permanent shocks (σ_ζ^2)	0.0267	
	(0.0135)	
Variance of transitory shocks (σ_ε^2)	0.0794	
	(0.0123)	
χ^2 goodness-of-fit statistic	15.21	15.22
	[0.0853]	[0.1241]

Note: The table reports simulated minimum χ^2 estimates of the parameters ϕ, λ, and α (asymptotic standard errors in parenthesis) and the χ^2 goodness-of-fit statistic (p-value of the test in square brackets). In column (2) we imposes the (acceptable) restriction that $\phi = 1$. The estimates of σ_ζ^2 and σ_ε^2 are obtained in a first step by imposing restrictions on the autocovariances of income growth (see Section 4).

model fits well the transition probabilities: We cannot reject the hypothesis that the empirical transition probabilities are jointly equal to the simulated ones.[24]

The simulation predicts almost perfectly the empirical transition matrix cell-by-cell, not just the aggregate mobility index. In Table 4 we report the simulated transition probabilities and (in parenthesis) the empirical transition probabilities, the same reported for 1993–1995 in Table 2. The comparison between the two sets

TABLE 4. Simulated and empirical transition matrix of consumption.

	1995 Quartile			
1993 Quartile	1st	2nd	3rd	4th
1st	0.6748	0.2515	0.0677	0.0061
	(0.6700)	(0.2528)	(0.0660)	(0.0112)
2nd	0.2513	0.4111	0.2748	0.0628
	(0.2416)	(0.4259)	(0.2665)	(0.0660)
3rd	0.0675	0.2764	0.4175	0.2386
	(0.0660)	(0.2653)	(0.4346)	(0.2341)
4th	0.0061	0.0613	0.2401	0.6926
	(0.0237)	(0.0549)	(0.2332)	(0.6883)

Note: The table reports the simulated consumption transitions between 1993 and 1995 and, in parenthesis, the empirical consumption transitions. The simulated transitions probabilities are obtained from the estimates reported in column (2) of Table 3.

24. Results for other years are similar with the exception of 1991–1993. In that period actual mobility increases substantially, a fact that is not captured by our simulations. One possible explanation is that the variance of the permanent shock, which is assumed to be time stationary, changed in 1993 due to the unprecedented strong recession. However, we cannot rule out that in 1993 the amount of measurement error is greater than in the other 2 years. Another possibility is that the 1993 recession impacted unevenly on households, a particular form of nonstationarity that we neglect in our simulation exercise.

of numbers is striking: Regardless of cell, the difference between the empirical and simulated values is at most 2 percentage points.

The estimated value of the excess sensitivity parameter ($\lambda = 0.16$) is broadly consistent with previous evidence on the effect of transitory income shocks on consumption expenditure. Using CEX quarterly panel data, Souleles (1999) and Parker (1999) examine, respectively, the response of household consumption to income tax refunds and to predictable changes in Social Security withholdings. Souleles finds evidence that the marginal propensity to consume is at least 35% of refunds within a quarter, and Parker that consumption reacts significantly to changes in tax rates. In both studies, the impact of transitory income shocks is too high to be consistent with the PIH model, but in the range of estimates produced by our hybrid model. Browning and Crossley (2001) survey several other studies reporting evidence that consumption overreacts to anticipated income innovations.

Our estimates allow us to characterize consumption mobility both in the short and in the long run by using recursively the transition law for consumption and the realizations of the income shocks (equation (15)). To illustrate, let us consider an individual who is in the bottom consumption quartile in the initial period. With full consumption insurance, absent measurement error and taste shocks, there is no mobility across quartiles: The individual rank in the consumption distribution is forever unchanged. Figure 1 shows that in a world in which households change consumption one-for-one in response to permanent income shocks, and smooth transitory shocks by saving and dissaving (the PIH model, obtained by setting $\phi = 1, \alpha = 0$, and $\lambda = 0$ in equation (16)), there is instead a 24% probability of leaving the bottom consumption quartile in period $t + 2$ conditional on being in the bottom quartile in period t. Because in each period the household receives new income shocks, we can generate a consumption distribution also for years

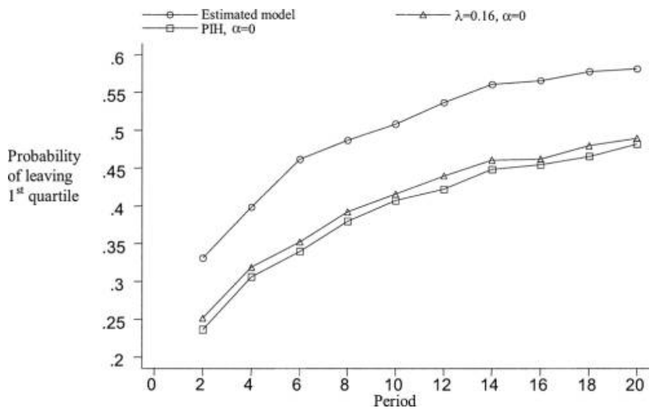

FIGURE 1. The probability of leaving the first quartile of the consumption distribution.

$t + 4, t + 6$, and so on until $t + 20$ (recall that our panel and transition law for consumption span 2 years of data). From each distribution we then create consumption quartiles and compute the probability of moving to higher quartiles in period $t+4$, $t+6$, and so on conditional on being in the first quartile in period t. This set of calculations traces the lowest line in Figure 1. Because the income process is nonstationary, income shocks compound and the chance of leaving the bottom quartile increases over time, up to a long-run value of 43%.

A second source of consumption mobility is the sensitivity of consumption to transitory income shocks. The intermediate line in Figure 1 is obtained using a transition law for consumption with $\phi = 1, \lambda = 0.16$, and $\alpha = 0$. Although the line lies above the one estimated for the PIH, the distance between the two is rather small, reflecting a small estimate for λ. Measurement error and taste shocks represent a third source of consumption mobility. The upper line in Figure 1 plots the estimated probability of moving from the lowest quartile for the full model ($\phi = 1, \lambda = 0.16$, and $\alpha = 0.38$, as in Table 3). The distance between this line and the intermediate line (with $\alpha = 0$) indicates that this composite term adds about 10 percentage points to the probability of leaving the bottom consumption quartile. Notice also that the probability in $t+2$ is 33%, matching the actual value ($0.33 = 1 - p_{11}$ in Table 5) and that measurement error impacts equally short- and long-run mobility.

5.2. Simultaneous Estimation of Consumption and Income Mobility

So far, we have followed the two-step estimation strategy described in Section 4. The first step uses covariance restrictions of the dynamic income process to estimate σ_ζ^2 and σ_ε^2, the variance of permanent and transitory income shocks. In the second step we estimate the parameters of the consumption rule, conditional

TABLE 5. Parameter estimates and χ^2 statistics.

ϕ	0.9776
	(0.0551)
λ	0.1542
	(0.0735)
α	0.3960
	(0.0192)
Variance of permanent shock (σ_ζ^2)	0.0246
	(0.0039)
Variance of transitory shock (σ_ε^2)	0.0501
	(0.0055)
χ^2 goodness-of-fit statistic	97.81
	[<0.0001]

Note: The table reports simulated minimum χ^2 estimates of the parameters $\phi, \lambda, \alpha, \sigma_\zeta^2$, and σ_ε^2 (asymptotic standard errors in parenthesis), and the χ^2 goodness-of-fit statistic (p-value of the test in square brackets).

on the estimates of σ_ζ^2 and σ_ε^2 obtained in the first step. Under our maintained assumptions this procedure is consistent but not fully efficient. For example, we do not use information of the income transition matrix—which may be interesting in its own right—or information on the joint behavior of consumption and income growth, if not through the dynamic consumption rule embedded in (15).

Here we extend our procedure and estimate the parameters of the income and consumption process simultaneously. We consider the restrictions that the theory imposes on the consumption transition matrix, the income transition matrix, and the Spearman joint rank correlation of income and consumption growth.[25] The estimation method is an extension of the one described in Appendix A, and it is detailed in Appendix C.

Table 5 reports the results of this exercise. The estimated variance of permanent income shocks, σ_ζ^2, is similar to the one estimated with the two-step procedure; the variance of transitory shocks, σ_ε^2, is slightly smaller (0.050 vs. 0.079) and more precisely estimated. The parameters of the consumption process ϕ, λ, and α are remarkably close to those reported in Table 3 and therefore yield a similar interpretation.

The goodness-of-fit statistics increases (there are now 20 degrees of freedom, instead of 9), which signals that the model could be improved. This can be seen from Table 6, where we report consumption and income transition probabilities, as well as the Spearman joint rank correlation index, predicted by our estimation strategy. The figures in parenthesis are the corresponding statistics computed from the actual data. Our model fits quite well the Spearman joint rank correlation, which is 0.42. As before, we are also able to fit consumption transition probabilities remarkably well given our estimated parameters. In contrast, the fit of the income transition matrix is not as good. For example, the theoretical model suggests that the probability of remaining in the third quartile of the income distribution should be 41%. In practice, in the data there is more persistence (48%). We leave investigations of the reasons for these discrepancies to future research, and in the remainder of the paper focus on the two-step procedure.

25. The Spearman index of the joint rank correlation between two variables x and y is defined as

$$R = \frac{\sum_{i=1}^{n} r(x_i) r(y_i) - n\left(\frac{n+1}{2}\right)^2}{\sqrt{\left[\sum_{i=1}^{n} r(x_i)^2 - n\left(\frac{n+1}{2}\right)^2\right]\left[\sum_{i=1}^{n} r(y_i)^2 - n\left(\frac{n+1}{2}\right)^2\right]}}$$

where $r(\cdot)$ is the rank and n the sample size.

The Spearman joint rank correlation index is distribution-free and is not affected by influential values. In terms of our competing theories, and in the absence of measurement error and taste shocks, a full consumption insurance model would suggest an index close to zero; a rule-of-thumb model would suggest an index close to 1. A hybrid model suggests values between these two extremes.

TABLE 6. Simulated and empirical transition matrices of consumption and income.

Panel (A): Consumption				
	1995 Quartile			
1993 Quartile	1st	2nd	3rd	4th
1st	0.6764	0.2502	0.0677	0.0056
	(0.6700)	(0.2528)	(0.0660)	(0.0112)
2nd	0.2506	0.4132	0.2739	0.0623
	(0.2416)	(0.4259)	(0.2665)	(0.0660)
3rd	0.0667	0.2752	0.4184	0.2398
	(0.0660)	(0.2653)	(0.4346)	(0.2341)
4th	0.0059	0.0617	0.2401	0.6923
	(0.0237)	(0.0549)	(0.2332)	(0.6883)
Panel (B): Income				
	1995 Quartile			
1993 Quartile	1st	2nd	3rd	4th
1st	0.6793	0.2460	0.0678	0.0068
	(0.6588)	(0.2204)	(0.0872)	(0.0336)
2nd	0.2436	0.4091	0.2766	0.0707
	(0.2503)	(0.4496)	(0.2204)	(0.0797)
3rd	0.0696	0.2785	0.4081	0.2438
	(0.0560)	(0.2565)	(0.4832)	(0.2042)
4th	0.0070	0.0667	0.2475	0.6787
	(0.0349)	(0.0736)	(0.2095)	(0.6820)
Panel (C): Joint rank correlation				
		0.4225		
		(0.4216)		

Note: Panel (A) of the table reports the simulated consumption transitions between 1993 and 1995 and, in parenthesis, the empirical consumption transitions. Panel (B) repeats for income. Panel (C) contains the Spearman joint rank correlation of consumption and income growth.

5.3. Measurement Error in Income

In this section we consider the robustness of our conclusions in the presence of measurement error in income. This error inflates the variance of the transitory shock but does not affect the variance of the permanent shock. To see this point, assume that true income is measured with a multiplicative error: $\ln y^*_{h,t} = \ln y_{h,t} + \nu_{h,t}$, where $\nu_{h,t}$ is an independently and identically normally distributed measurement error with mean zero and variance σ^2_ν. Using the income process (4)–(5):

$$\ln y^*_{h,t} = \beta X_{h,t} + p_{h,t} + \varepsilon_{h,t} + \nu_{h,t},$$

the 2-years-apart income residual is now

$$u_{h,t} - u_{h,t-2} = \zeta_{h,t} + \zeta_{h,t-1} + \varepsilon_{h,t} - \varepsilon_{h,t-2} + \nu_{h,t} - \nu_{h,t-2}.$$

The covariance matrix of the first difference of the income residuals depends now on the variance of the measurement error:

$$E[(u_{h,\tau} - u_{h,\tau-2})^2] = 2\sigma_\zeta^2 + 2\sigma_\varepsilon^2 + 2\sigma_\nu^2$$
$$E[(u_{h,\tau} - u_{h,\tau-2})(u_{h,\tau-2} - u_{h,\tau-4})] = -\sigma_\varepsilon^2 - \sigma_\nu^2$$
$$E[(u_{h,\tau} - u_{h,\tau-2})(u_{h,\tau-j} - u_{h,\tau-j-2})] = 0 \text{ for all } j \geq 4.$$

However, it can be checked that measurement error inflates the estimated variance of the transitory shock by σ_ν^2, but not the variance of the permanent shock σ_ζ^2, which is still identified by the difference between the variance and (minus twice) the first-order autocovariance. The conclusion is that even though the estimate of the variance of the permanent shock is unaffected by serially uncorrelated measurement error, the estimate of the variance of the transitory shock is not.

This implies that in the model with full consumption insurance, idiosyncratic income shocks play no role regardless of measurement error in income. In the permanent income model, the impact of measurement error in income is bound to be small, because transitory shocks play a very limited role. In contrast, measurement error may have a large impact in the rule-of-thumb model. Because we cannot identify σ_ν^2 from the data, we repeat our simulation: (a) dropping the self-employed from the sample on which we estimate the income process[26] and (b) downsizing the variance of the transitory shock, that is, assuming that one-third or one-half of the estimated first-order autocovariance reflects measurement error. The results of these experiments are very similar to the simulations reported in Tables 3 and 4 and are not reported for brevity.

5.4. Group Estimates

Except for the extreme case of consumption insurance, models with incomplete insurance suggest that if different population groups are systematically exposed to different idiosyncratic shocks (and therefore face different income processes), consumption mobility should differ across groups in a predictable way.[27] Therefore, comparison of different population groups with different income-generating processes is potentially quite interesting. Indeed, even more compelling evidence for the ability of our simulations to explain consumption transitions comes from comparing consumption mobility in two education groups: compulsory schooling or less and high school or college degree.

26. Brandolini and Cannari (1994) note that in the SHIW income from self-employment is less well estimated than wages or salaries.

27. Attanasio and Davis (1996) also exploit predictable differences between education groups to provide insights about the consumption insurance hypothesis.

Focus on education is warranted for at least three reasons: (1) education is an exogenous characteristic by which one can partition the sample; (2) there is wide evidence that different education groups face different earnings opportunities and uncertainties; and (3) education is likely to be correlated with variables affecting preferences and therefore with different consumption behavior. We run the income regressions separately for households headed by individuals with high and low education. We then estimate the autocovariance matrix as explained in Section 4, and find $\sigma_\zeta^2 = 0.0296$ and $\sigma_\varepsilon^2 = 0.0754$ for the less well-educated, and $\sigma_\zeta^2 = 0.0198$ and $\sigma_\varepsilon^2 = 0.0895$ for those with at least a high school degree. The estimated variances signal that the less well-educated face a higher variance of permanent income shocks, a pattern also uncovered by Carrol and Samwick (1997) with U.S. data. Because in our sample the income process varies considerably by education groups, we have an ideal setting to test the validity of models of intertemporal choice and of the robustness of our procedure.

The transition probabilities reported in the two lower panels of Table 7 indicate also that consumption mobility differs between the two groups in a systematic way. Applying the test on difference of transition probabilities outlined in Section 2, we reject the hypothesis that the two matrices are equal at the 1% significance level.

Quite clearly, the consumption insurance model is unable to explain differences in consumption mobility emerging from income shocks, transitory or permanent. In that model all shocks are insured, and provided measurement error or taste shocks are the same in both groups, consumption mobility between two groups exposed to different shocks should be identical. Therefore, the fact that mobility is higher in the group with lower education provides further evidence against the consumption insurance model.[28] For quite different reasons, the rule-of-thumb model with $\lambda = 1$ (or any model where excess sensitivity to transitory income shocks plays a prominent role) predicts little or no difference between education groups. In the simulations the lower variance of the transitory shock for the less well-educated is offset by a higher variance of the permanent shock, resulting in approximately the same mobility rates in the two groups.

We therefore estimate equation (15) and the associated consumption transitions allowing for differential response between the two education groups. The parameter estimates of the simulated minimum χ^2 method and the associated χ^2 statistic are reported in Table 7. Also in this case, we cannot reject the hypothesis that $\phi = 1$ in each of the two groups. We find a value of $\alpha = 0.38$ (s.e. 0.01) and $\lambda = 0.4$ (0.05) in the group with low education and $\alpha = 0.28$ (0.01) and $\lambda = 0.09$ (0.05) in the group with high school or college degree. The model replicates quite

28. The different estimates of α in the two groups will generate some differences in estimated mobility. However, unlike the estimates of λ, the α estimates are fairly similar in the two groups, and so this is unlikely to be a relevant explanation.

TABLE 7. Simulated and empirical mobility for different education groups.

Panel(A) Summary statistics		
	Low education	High education
Variance of permanent shock (σ_ζ^2)	0.0296	0.0198
	(0.0115)	(0.0347)
Variance of transitory shock (σ_ε^2)	0.0754	0.0895
	(0.0097)	(0.0329)
λ	0.3991	0.0889
	(0.0459)	(0.0470)
α	0.3814	0.2846
	(0.0120)	(0.0125)
χ^2 goodness-of-fit statistic	23.03	20.85
	[0.0106]	[0.0222]
χ^2 test of parameter equality	55.85	
	[7.4e–013]	

Panel(B): Probabilities Low education				
	1995 Quartile			
1993 Quartile	1st	2nd	3rd	4th
1st	0.6432	0.2610	0.0845	0.0113
	(0.6449)	(0.2633)	(0.0830)	(0.0088)
2nd	0.2413	0.3863	0.2848	0.0877
	(0.2240)	(0.4208)	(0.2532)	(0.1020)
3rd	0.0733	0.2654	0.3913	0.2700
	(0.0780)	(0.2524)	(0.4421)	(0.2277)
4th	0.0099	0.0770	0.2498	0.6632
	(0.0327)	(0.0750)	(0.2500)	(0.6423)

Panel(C): Probabilities High education				
	1995 Quartile			
1993 Quartile	1st	2nd	3rd	4th
1st	0.7037	0.2485	0.0457	0.0022
	(0.7016)	(0.2258)	(0.0605)	(0.0121)
2nd	0.2472	0.4463	0.2686	0.0379
	(0.2308)	(0.4656)	(0.2389)	(0.0648)
3rd	0.0464	0.2681	0.4661	0.2194
	(0.0242)	(0.2460)	(0.4839)	(0.2460)
4th	0.0015	0.0373	0.2156	0.7456
	(0.0163)	(0.0569)	(0.2398)	(0.6870)

Note: The first panel reports summary statistics for two education groups: the variance of the income shocks, the estimates of α and λ (asymptotic standard errors in parenthesis), the associated χ^2 goodness-of-fit statistic (p-value of the test in square brackets), and the χ^2 statistic of the test that the parameters of the two education groups are the same (p-value in square brackets). The education groups are defined as compulsory schooling or less, and high school or college. The other two panels report the simulated transition probabilities and (in parenthesis) the empirical transition probabilities for the two education groups.

well also the difference in empirical and simulated mobility between the two groups: The simulated mobility index (0.64 and 0.55 for low and high education, respectively) is quite close to empirical mobility in each group. And in each of the two cases the χ^2 statistic does not reject the null hypothesis that the simulated probabilities are equal to the empirical ones at the 1% significance level.

As a final check of the validity of the estimates, we test whether the parameters are the same in the two groups. Call $\tilde{\theta}_h$ and $\tilde{\theta}_l$ the $k \times 1$ vectors of simulated minimum χ^2 estimates of θ for high- and low-educated individuals. Given the asymptotic normal distribution of the estimator and the fact that the two samples are independent, the null hypothesis of no group difference can be tested using the statistic

$$(\tilde{\theta}_h - \tilde{\theta}_l)'(\mathrm{var}(\tilde{\theta}_h) + \mathrm{var}(\tilde{\theta}_l))^{-1}(\tilde{\theta}_h - \tilde{\theta}_l),$$

which is distributed χ^2_k under the null. The test statistic, reported in the last row of the first panel of Table 7, rejects the null hypothesis of parameter equality. The other two panels of Table 7 report the simulated transition probabilities and (in parenthesis) the empirical transition probabilities for the two education groups. Once more, each of the simulated probabilities is remarkably close to the empirical transitions regardless of the group considered.

From an economic point of view, the result that the less-well-educated individuals are more responsive to transitory income shocks than the high-income group is of particular interest. To the extent that these households are less likely to have access to credit and insurance markets than households with higher education, our findings support the hypothesis that excess sensitivity stems from the effect of borrowing constraints, rather than from other sources.[29] We also find that the estimate of α is higher for the less well-educated. Whether this reflects a tendency to report noisier consumption data or a greater incidence of taste shocks in this group is, however, an issue that cannot be settled here.

5.5. Relation with Previous Tests

It is useful to contrast our approach with previous tests of models of intertemporal choice. First of all, our simulation method produces estimates of the propensity to consume out of transitory and permanent income shocks. These parameters are of great policy interest, for instance to evaluate the effect of a tax cut or other changes in the household budget constraint. Excess sensitivity of consumption has sometimes been inferred from the income growth coefficient in Euler equations estimates. However, there is much disagreement concerning the interpretation of the excess sensitivity parameter due to various identification problems in the estimation of the Euler equation (Attanasio 2000). Whereas the Euler equation

29. Jappelli, Pischke, and Souleles (1998) report evidence from the Survey of Consumer Finances that individuals with less than a college degree are more likely to be turned down for loans, to have no credit card, or to have no line of credit. The same households have also fewer assets relative to income, an indicator that has often been interpreted as bearing on the incidence of borrowing constraints.

literature is concerned with estimation of preference parameters derived from the first-order conditions of the consumers' optimization problem, we attempt at estimating the parameters of the consumption rule. This does not come without costs, however. We make specific assumptions about preferences and the income generating process, and our estimates are therefore conditional on the validity of the theoretical framework and on the stability of the income process. This paper is therefore part of a growing literature in macroeconomics that attempts to estimate structural (or semistructural) models by means of simulated estimation methods.

Second, and for quite different reasons, our approach to test for consumption insurance differs from previous tests based on regression analysis. Cochrane (1991), Mace (1991), Townsend (1994), Attanasio and Davis (1996), and Zhang and Ogaki (2001) regress household consumption growth on aggregate variables and idiosyncratic shocks (such as change in disposable income, unemployment hours, and days of illness). The implication of the theory is that none of these shocks should impact household consumption growth, as in equation (9). Focussing instead on the relation between consumption insurance and consumption mobility has two advantages: (1) we do not need to identify explicitely any of these shocks, and (2) we don't need to assume that they are uncorrelated with unobservable or omitted preference shocks, including household fixed effects. Moreover, measurement error in the shock variables biases tests based on regression analysis towards the null hypothesis of full consumption insurance; our testing strategy is instead robust to such problem. Of course, as we have clarified, our approach requires us to make specific (and untestable) assumptions on how taste shocks impact the marginal utility of consumption.

Deaton and Paxson (1994) and Attanasio and Jappelli (2001) test another implication of the theory of consumption insurance, namely, that the cross-sectional variance of consumption is constant over time. However, the distribution of consumption at time t might have the same variance of the distribution at time $t-1$ even if there is mobility in the underlying distributions.[30] Tests based on the dynamics of the cross-sectional variance of consumption are therefore biased towards the null. Our test instead still signals rejection of the consumption insurance model even in situations in which the cross-sectional variance is constant over time but there is mobility in the underlying distribution.

6. Summary

The implications of the theories of intertemporal consumption choice for consumption mobility are as yet unexplored. In this paper we study transition

30. For instance, suppose that a poor and a rich household switch ranks in the consumption distribution. This will not change the cross-sectional variance of consumption but represents a violation of consumption insurance.

probabilities for total nondurable consumption using the 1987–1995 panel contained in the Bank of Italy Survey of Household Income and Wealth. The panel data allow us to calculate an empirical transition matrix of log per capita consumption. The matrix shows that there is substantial consumption mobility: In any year, about 60% of the households moves up or down in the consumption distribution.

In the remainder of the paper we attempt to understand which model of intertemporal consumption choice is capable of explaining the amount of consumption mobility we observe in the data. From the theoretical point of view, the consumption insurance model provides the clearest implications for consumption mobility. In a model where all idiosyncratic income shocks are insured, the initial cross-sectional distribution of consumption is a sufficient statistic for all future distributions, and therefore, apart from measurement error in consumption and taste shocks, the model predicts zero consumption mobility. On the other hand, the rule-of-thumb model is one where income shocks have the greatest impact on consumption; it therefore generates substantial consumption mobility. Finally, in models with optimizing agents and incomplete markets (such as the permanent income model or models with precautionary saving) households react mainly to permanent income shocks. Thus, the degree of mobility predicted by the model is intermediate between the two other models.

We carefully parametrize an income process to distinguish between transitory and permanent shocks and use the estimated parameters to simulate theoretically the degree of mobility stemming from each of the consumption models examined. We then compare them statistically with the actual amount of mobility estimated in the data. The results reject statistically each of the simple representations of the consumption decision rule, and reveal that households smooth income shocks to a lesser extent than implied by the PIH. A noteworthy feature of our method is that the estimates are robust to the presence of measurement error in consumption and taste shocks, although we cannot identify the separate contribution of these two components on consumption dynamics.

Several criteria suggest that our estimates describe the dynamics of the consumption distribution remarkably well. First, the estimates are able to match the empirical transition matrix cell by cell. Second, the results are robust with respect to different definitions of consumption (in per capita or per adult equivalent terms), to the presence of measurement error in income, taste shocks, and to various other sensitivity checks on sample exclusions and definitions. Third, the results do not change when we estimate income and consumption mobility simultaneously, instead of relying on a two-step procedure. Finally, and most important, the group-specific estimates by education match the different patterns of consumption mobility we find in the data.

There are three important by-products of our analysis. First, we produce estimates of the sensitivity of consumption to permanent and transitory income shocks that are potentially useful to evaluate fiscal policy experiments that affect the timing of income receipts and, more generally, households' budget constraints.

In this respect, we find considerable asymmetric response to transitory income shocks by education groups: a low response in the group with higher education and a relatively high response for households with lower education.

Second, we provide a powerful test of the consumption insurance model. So far these tests have focused on mean and variance restrictions of the distribution of consumption growth. Mean restrictions require consumption growth to be orthogonal, on average, to idiosyncratic income shocks. If shocks are measured with error, however, these tests are biased towards the null hypothesis of full consumption insurance. Variance restrictions require the cross-sectional variance of consumption growth to be constant over time. But the variance might be stationary even if the underlying consumption distribution is shifting. Thus, variance restriction tests too are biased towards the null. Our test is free from these problems, because we look at the entire consumption distribution, not just its mean or variance. On the other hand, the implementation of this test and, more generally, the evaluation of consumption mobility requires genuine panel data and suitable assumptions about the distribution of measurement error and taste shocks, while mean and variance restriction tests can be performed with repeated cross-sectional data.

Finally, the estimates could be used to single out the separate contributions of incomplete markets, excess sensitivity, measurement error, and taste shocks in generating the short and long run consumption mobility we observe in the panel data. One important question is to what extent the failure of complete markets is due to the unwillingness of society to forgo social mobility, an issue that we plan to explore in future research.

Appendix A: The Simulated Minimum χ^2 Estimator

Let P(θ) represent the $q \times q$ transition matrix with typical element $p_{ij}(\theta)$, where θ is a vector of k unknown parameters:

$$\mathbf{P}(\theta) = \begin{bmatrix} p_{11}(\theta) & p_{12}(\theta) & \dots & p_{1q}(\theta) \\ p_{21}(\theta) & p_{22}(\theta) & \dots & p_{2q}(\theta) \\ \dots & \dots & \dots & \dots \\ p_{q1}(\theta) & p_{q2}(\theta) & \dots & p_{qq}(\theta) \end{bmatrix}.$$

Conformably with $\mathbf{P}(\theta)$ let $\hat{\mathbf{P}}$ represent the $q \times q$ empirical transition matrix with typical element $\hat{p}_{ij}$:

$$\hat{\mathbf{P}} = \begin{bmatrix} \hat{p}_{11} & \hat{p}_{12} & \dots & \hat{p}_{1q} \\ \hat{p}_{21} & \hat{p}_{22} & \dots & \hat{p}_{2q} \\ \dots & \dots & \dots & \dots \\ \hat{p}_{q1} & \hat{p}_{q2} & \dots & \hat{p}_{qq} \end{bmatrix}.$$

The transition matrices $\mathbf{P}(\theta)$ and $\hat{\mathbf{P}}$ are subject to the restrictions $\sum_{j=1}^{q} p_{ij}(\theta) = 1$ and $\sum_{j=1}^{q} \hat{p}_{ij} = 1$ $(i = 1 \ldots q)$, respectively. This creates a singularity problem similar to the one in the estimation of a full demand system. To avoid this problem, we drop one column (say, the qth column) from both $\mathbf{P}(\theta)$ and $\hat{\mathbf{P}}$.

Let $\mathbf{p}(\theta)$ represent the $q(q-1) \times 1$ vector of true transition probabilities and conformably with $\mathbf{p}(\theta)$ let $\hat{\mathbf{p}}$ represent the $q(q-1) \times 1$ vector of estimated transition probabilities. The distance between the empirical and true transition probabilities is $\mathbf{d}(\theta) = \hat{\mathbf{p}} - \mathbf{p}(\theta)$, whose covariance matrix $\Omega(\theta)$ is block diagonal with generic block[31]

$$\Omega_i(\theta) = \begin{bmatrix} \frac{p_{i1}(\theta)(1-p_{i1}(\theta))}{n_i} & -\frac{p_{i1}(\theta)p_{i2}(\theta)}{n_i} & \ldots & -\frac{p_{i1}(\theta)p_{iq-1}(\theta)}{n_i} \\ & \frac{p_{i2}(\theta)(1-p_{i2}(\theta))}{n_i} & \ldots & -\frac{p_{i2}(\theta)p_{iq-1}(\theta)}{n_i} \\ & & \ldots & \ldots \\ & & & \frac{p_{iq-1}(\theta)(1-p_{iq-1}(\theta))}{n_i} \end{bmatrix}$$

for $i = 1 \ldots q$ (we assume $n_i = n/q$ is an integer for simplicity), so that

$$\Omega(\theta) = \begin{bmatrix} \Omega_1(\theta) & 0 & \ldots & 0 \\ & \Omega_2(\theta) & \ldots & 0 \\ & & \ldots & \ldots \\ & & & \Omega_q(\theta) \end{bmatrix}.$$

From Chamberlain (1984), the minimum χ^2 method solves the problem

$$\min_{\theta} \mathbf{d}(\theta)' \mathbf{W} \mathbf{d}(\theta),$$

where $\mathbf{W}$ is a weighting matrix. Call $\hat{\theta}$ the minimum χ^2 estimate of θ. Chamberlain (1984) and others show that $\hat{\theta}$ is consistent, asymptotically normal with covariance matrix

$$\operatorname{var}(\hat{\theta}) = \left(\mathrm{G}(\hat{\theta})'\mathrm{W}\mathrm{G}(\hat{\theta})\right)^{-1} \mathrm{G}(\hat{\theta})'\mathrm{W}\Omega(\theta)\mathrm{W}\mathrm{G}(\hat{\theta}) \left(\mathrm{G}(\hat{\theta})'\mathrm{W}\mathrm{G}(\hat{\theta})\right)^{-1},$$

where $G(\hat{\theta}) = \partial d(\hat{\theta})/\partial\theta'$ is the gradient matrix. It is a well-known result that the optimal weighting matrix (in the efficiency sense) is $\Omega(\theta)^{-1}$. In this case,

$$\operatorname{var}(\hat{\theta}) = \left(\mathrm{G}(\hat{\theta})'\Omega(\theta)^{-1}\mathrm{G}(\hat{\theta})\right)^{-1}.$$

In our case $\mathbf{p}(\theta)$ has no closed form, so we replace it with an approximation based on simulations, as in the simulated method of moments (McFadden 1989; Duffie and Singleton 1991). Recall that the generic $p_{ij}(\theta)$ is

$$p_{ij}(\theta) = \Pr(\ln c_{h,t}^{*} \in i \,|\, \ln c_{h,t-2}^{*} \in j, \theta),$$

31. We neglect the extra randomness induced by the fact that the class boundaries are preestimated.

for example, the probability of making a transition to class i from class j conditioning on being in class j. The transition law for consumption is determined by (15), reproduced here:

$$\ln c^*_{h,t} = \ln c^*_{h,t-2} + \phi\left(\frac{\lambda + r}{1+r}\varepsilon_{h,t} + \frac{(1-\lambda)r}{1+r}\varepsilon_{h,t-1} - \lambda\varepsilon_{h,t-2} + \zeta_{h,t} + \zeta_{h,t-1}\right) + \alpha\big(\upsilon_{h,t} - \upsilon_{h,t-2}\big),$$

where in our case $\theta = (\varphi\ \alpha\ \lambda)'$.[32] For simulation purposes, we assume $\varepsilon_{h,r} \sim N\big(0, \sigma^2_\varepsilon\big)$, $\zeta_{h,\tau} \sim N\big(0, \sigma^2_\zeta\big)$, and $\upsilon_{h,\tau} \sim N\big(0, \sigma^2_{\ln c^*_{h,\tau}}\big)$ for all τ.

By construction, the normality of the income shocks and of measurement error generates a symmetric transition matrix for consumption. This feature of the simulations is consistent with the symmetry of the empirical matrix documented in Table 2. Our results do not depend on the normality assumption. We choose normality for simplicity, but note that any symmetric distribution would work as well, because it would imply a symmetric transition matrix.

Define $\mathbf{u}_h = (\varepsilon_{h,t}\ \varepsilon_{h,t-1}\ \varepsilon_{h,t-2}\ \zeta_{h,t}\zeta_{h,t-1}\ \upsilon_{h,t}\ \upsilon_{h,t-2})'$ the vector of disturbances. For each household h, we draw S independent realizations of $\mathbf{u}_h$, and store the HS realizations (H being the number of households).[33] It is necessary to keep these basic drawings of $\mathbf{u}^s_h$ fixed when θ changes, in order to have good numerical and statistical properties of the estimators based on the simulations.

Conditioning on the measured (not simulated) $\ln c^*_{h,t-2}$, the simulated $\mathbf{u}^s_h$, and a choice for θ, one obtains

$$\ln c^{*s}_{h,t} = \ln c^*_{h,t-2} + \phi\left(\frac{\lambda + r}{1+r}\varepsilon^s_{h,t} + \frac{(1-\lambda)r}{1+r}\varepsilon^s_{h,t-1} - \lambda\varepsilon^s_{h,t-2} + \zeta^s_{h,t} + \zeta^s_{h,t-1}\right) + \alpha\big(\tilde{\upsilon}^s_{h,t} - \tilde{\upsilon}^s_{h,t-2}\big), \tag{A.1}$$

This allows computation of $p^s_{ij}(\theta)$. One can then define $\bar{p}_{ij}(\theta) = S^{-1}\sum_{s=1}^{S} p^s_{ij}(\theta)$ as the approximation of $p_{ij}(\theta)$ obtained by means of simulations.

Call the simulated distance $\bar{\mathbf{d}}(\theta) = \hat{\mathbf{p}} - \bar{\mathbf{p}}(\theta)$ where $\bar{\mathbf{p}}(\theta)$ is the vector of simulated transition probabilities with generic element $\bar{p}_{ij}(\theta)$. Note that the covariance matrix of $\bar{\mathbf{d}}(\theta)$, $\bar{\Omega}(\theta) \overset{\text{a.s.}}{\to} (1 + S^{-1})\,\Omega\,(\theta)$ where $(1 + S^{-1})$ is an inflating factor of the variance of the true distance vector induced by the additional randomness of the simulations. With a large enough number of simulations, however, this factor plays little weight in practice.

The choice of θ minimizes the simulated minimum χ^2 criterion

$$\min_\theta \bar{\mathbf{d}}(\theta)'\bar{\Omega}(\theta)^{-1}\bar{\mathbf{d}}(\theta).$$

32. We neglect the problems associated with the fact that r is given, and that $\sigma^2_{\ln c^*_{h,t-2}}$, $\sigma^2_{\ln c^*_{h,t}}$, σ^2_ε, and σ^2_ζ are preestimated.

33. In each year we choose a sample size identical to the number of actual sample transitions (for instance, it is 2,982 in 1991–1993 and 3,211 in 1993–1995).

Call $\tilde{\theta}$ the resulting solution. Then, the results in Lee and Ingram (1991), McFadden (1989), and Duffie and Singleton (1993) imply that $\tilde{\theta}$ is consistent, asymptotically normal with covariance matrix

$$\operatorname{var}(\tilde{\theta}) = \left(1 + \frac{1}{S}\right)\left[\mathrm{G}(\tilde{\theta})'\,\Omega(\tilde{\theta})^{-1}\,\mathrm{G}(\tilde{\theta})\right]^{-1}.$$

Goodness of fit can be assessed using

$$m = \mathbf{d}(\tilde{\theta})'\,\Omega(\tilde{\theta})^{-1}\,\mathbf{d}(\tilde{\theta}) \sim \chi^2_{q(q-1)-k}.$$

Note that when $q(q-1) > k$ as in our empirical application, this goodness-of-fit statistic can be interpreteted as an overidentifying restriction statistics. This is because we estimate k parameters but minimize the distance between $q(q-1) > k$ actual and theoretical transition probabilities.

The algorithm that we implement is thus the following:

1. Draw $\mathbf{u}_h^s$ ($h = 1 \ldots H, s = 1 \ldots S$).
2. Choose a starting value for θ, say θ_0.
3. Compute $\ln c^*_{h,t}$ using (A1), $\bar{p}_{ij}(\theta_0)$, and $\bar{d}_{ij}(\theta_0) = \hat{p}_{ij} - \bar{p}_{ij}(\theta_0)$ ($i = 1 \ldots q$, $j = 1 \ldots q-1$).
4. Compute $\bar{\mathbf{d}}(\theta_0)'\,\bar{\Omega}(\theta_0)^{-1}\bar{\mathbf{d}}(\theta_0)$.
5. Update the value of θ.
6. Repeat steps 3–5 until a prespecified convergence criterion is met. Eventually this provides the required simulated minimum χ^2 estimate $\tilde{\theta}$ of θ.

We update the value of θ using the simulated annealing method of Goffe, Ferrier, and Rogers (1994).[34] This is a derivative-free minimization method that escapes local minima. Starting from an initial value, the algorithm takes a step and evaluates the function. Downhill steps are always accepted, whereas uphill steps are accepted probabilistically according to the Metropolis criterion. As the algorithm proceeds, the length of the step declines until the χ^2 reaches the global minimum.

Appendix B: Test Equivalence

Here we prove the statement in Section 2 that the χ^2 goodness-of-fit criterion (2)

$$\sum_{i=1}^{q}\sum_{j=1}^{q} n_i \frac{\left(\hat{p}_{ij} - p_{ij}(\theta)\right)^2}{p_{ij}(\theta)} \tag{A.2}$$

is equivalent to $(\hat{\mathbf{p}} - \mathbf{p}(\theta))\,\Omega(\theta)^{-1}(\hat{\mathbf{p}} - \mathbf{p}(\theta))'$.

34. We use the Gauss code on simulated annealing written by E.G. Tsionas and available at ⟨http://www.american.edu/academic.depts/cas/econ/gaussres/optimize/optimize.htm⟩.

Note first that (A2) is the sum of q independent χ^2 distributions of the form $m_i = \sum_{j=1}^{q} n_i p_{ij}^{-1}(\hat{p}_{ij} - p_{ij}(\theta))^2$. The sum of q independent χ^2 distributions is also a χ^2 distribution with degrees of freedom equal to the sum of the degrees of freedom of the χ^2 distributions that are summed.

Notice that the theoretical and empirical transition probabilities are subject to the restrictions $\sum_{j=1}^{q} p_{ij}(\theta) = 1$ and $\sum_{j=1}^{q} \hat{p}_{ij} = 1 (i = 1 \ldots q)$, respectively. Thus $m_i = \sum_{j=1}^{q} n_i p_{ij}(\theta)^{-1}(\hat{p}_{ij} - p_{ij}(\theta))^2$ can be rewritten as

$$\begin{aligned} m_i &= \sum_{j=1}^{q-1} \frac{n_i}{p_{ij}(\theta)}(\hat{p}_{ij} - p_{ij}(\theta))^2 + \frac{n_i}{p_{iq}(\theta)} \left[\sum_{j=1}^{q-1} \left(\hat{p}_{ij} - p_{ij}(\theta)\right) \right]^2 \\ &= \sum_{j=1}^{q-1} \frac{n_i}{p_{ij}(\theta)}\left(\hat{p}_{ij} - p_{ij}(\theta)\right)^2 + \frac{n_i}{p_{iq}(\theta)} \sum_{j=1}^{q-1} (\hat{p}_{ij} - p_{ij}(\theta)) \sum_{j=1}^{q-1} (\hat{p}_{ij} - p_{ij}(\theta)), \end{aligned}$$

or, more compactly

$$\begin{aligned} m_i &= \mathbf{d}_i(\theta)'\mathbf{A}_i(\theta)\, \mathbf{d}_i(\theta) + \mathbf{d}_i(\theta)'\mathbf{B}_i(\theta)\, \mathbf{d}_i(\theta) \\ &= \mathbf{d}_i(\theta)'\Lambda_i(\theta)\, \mathbf{d}_i(\theta), \end{aligned}$$

where $\mathbf{d}_i(\theta) = \hat{\mathbf{p}}_i - \mathbf{p}_i(\theta)$ is the distance between empirical and true transition probabilities in row i of the transition matrix (excluding the qth column), and $\Lambda_i(\theta) = \mathbf{A}_i(\theta) + \mathbf{B}_i(\theta)$, with

$$\mathbf{A}_i(\theta) = \begin{bmatrix} \frac{n_i}{p_{11}(\theta)} & 0 & \ldots & 0 \\ 0 & \frac{n_i}{p_{12}(\theta)} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \frac{n_i}{p_{1q-1}(\theta)} \end{bmatrix}$$

and $\mathbf{B}_i(\theta) = (n_i/p_{iq}(\theta))\, \mathrm{ii}'$, where $\mathbf{i}$ is a $(q-1) \times 1$ vector of ones, so that $\mathbf{B}_i(\theta)$ is a matrix that contains $(n_i/p_{iq}(\theta))$ everywhere. It's easy to prove that $\Lambda_i(\theta) = \Omega_i(\theta)^{-1}$ defined in Appendix B. Because asymptotically $\mathrm{d}_i(\theta) \sim N(0, \Omega_i(\theta))$, it follows that

$$m_i = \sum_{j=1}^{q} n_i \frac{(\hat{p}_{ij} - p_{ij}(\theta))^2}{p_{ij}(\theta)} = \mathbf{d}_i(\theta)'\Omega_i(\theta)^{-1}\mathbf{d}_i(\theta)$$

is distributed χ^2 with $(q-1)$ degrees of freedom. Moreover

$$\begin{aligned} \sum_{i=1}^{q}\sum_{j=1}^{q} n_i \frac{(\hat{p}_{ij} - p_{ij}(\theta))^2}{p_{ij}(\theta)} &= \sum_{i=1}^{q} m_i = \sum_{i=1}^{q} \mathbf{d}_i(\theta)'\Omega_i(\theta)^{-1}\mathbf{d}_i(\theta) \\ &= \mathbf{d}(\theta)'\Omega(\theta)^{-1}\mathbf{d}(\theta) \end{aligned}$$

is distributed χ^2 with $q(q-1)$ degrees of freedom. This is exactly the function that we minimize in the simulated minimum χ^2 application. This proves the equivalence between $\sum_{i=1}^{q}\sum_{j=1}^{q} n_i p_{ij}(\theta)^{-1}(\hat{p}_{ij} - p_{ij}(\theta))^2$ and $\mathbf{d}(\theta)'\Omega(\theta)^{-1}\mathbf{d}(\theta)$.

An alternative to the minimum χ^2 criterion $\sum_{i=1}^{q}\sum_{j=1}^{q} n_i p_{ij}(\theta)^{-1}$ $(\hat{p}_{ij} - p_{ij}(\theta))^2$ is to use the modified minimum χ^2 criterion $\sum_{i=1}^{q}\sum_{j=1}^{q} n_i \hat{p}_{ij}^{-1}$ $(\hat{p}_{ij} - p_{ij}(\theta))^2$. Following the same steps above, one can show that

$$\sum_{i=1}^{q}\sum_{j=1}^{q} n_i \frac{(\hat{p}_{ij} - p_{ij}(\theta))^2}{\hat{p}_{ij}} = \mathbf{d}(\theta)'\hat{\Omega}^{-1}\mathbf{d}(\theta),$$

where $\hat{\Omega}$ is a block-diagonal matrix with generic block

$$\hat{\Omega}_i = \begin{bmatrix} \frac{\hat{p}_{i1}(1-\hat{p}_{i1})}{n_i} & -\frac{\hat{p}_{i1}\hat{p}_{i2}}{n_i} & \cdots & -\frac{\hat{p}_{i1}\hat{p}_{iq-1}}{n_i} \\ & \frac{\hat{p}_{i2}(1-\hat{p}_{i2})}{n_i} & \cdots & -\frac{\hat{p}_{i2}\hat{p}_{iq-1}}{n_i} \\ & & \cdots & \cdots \\ & & & \frac{\hat{p}_{iq-1}(1-\hat{p}_{iq-1})}{n_i} \end{bmatrix}.$$

Because $\hat{p}_{ij}$ is a consistent estimate of $p_{ij}(\theta)$, $\hat{\Omega} \to^{\text{a.s.}} \Omega(\theta)$. In the estimation, we use the modified simulated minimum χ^2 criterion, that is, use $\hat{\Omega}$ (based on the empirical transition probabilities) as an estimate of $\Omega(\theta)$.

Appendix C: The Simulated Minimum χ^2 Estimator in the Extended Case

Let $\mathbf{P}^x(\theta)$ be a $q \times q$ transition matrix for variable x with typical element $\mathrm{p}_{ij}^x(\theta)$; $\hat{\mathbf{P}}^x$ is its empirical analog with typical element $\hat{P}_{ij}^x$. Let $\mathbf{p}^x(\theta)$ be the stacked $q(q-1)\times 1$ vector of true transition probabilities obtained after dropping the qth column of the transition matrix $\mathbf{P}^x(\theta)$ to avoid singularity; and $\hat{\mathbf{p}}^x$ its empirical analog.

Define

$$\mathbf{p}(\theta) = \begin{pmatrix} \mathbf{p}^c(\theta) \\ \mathbf{p}^y(\theta) \\ R(\theta) \end{pmatrix} \text{ and } \hat{\mathbf{p}} = \begin{pmatrix} \hat{\mathbf{p}}^c \\ \hat{\mathbf{p}}^y \\ \hat{R} \end{pmatrix},$$

where superscripts c and y refer to (log) consumption and income, respectively, $R(\theta)$ is the Spearman joint rank correlation of true consumption and income growth, and $\hat{R}$ its empirical analog.

The distance between empirical and true statistics is $\mathbf{d}(\theta) = \hat{\mathbf{p}} - \mathbf{p}(\theta)$, whose covariance matrix $\Omega(\theta)$ we assume to be block-diagonal[35]

$$\Omega(\theta) = \begin{bmatrix} \Omega^c(\theta) & 0 & 0 \\ & \Omega^y(\theta) & 0 \\ & & \omega(\theta) \end{bmatrix},$$

with $\omega(\theta) = (1 - R(\theta)^2)/(n-2)$,

$$\Omega^x(\theta) = \begin{bmatrix} \Omega_1^x(\theta) & 0 & \dots & 0 \\ & \Omega_2^x(\theta) & \dots & 0 \\ & & \dots & \dots \\ & & & \Omega_q^x(\theta) \end{bmatrix},$$

and

$$\Omega_i^x(\theta) = \begin{bmatrix} \frac{p_{i1}^x(\theta)\left(1-p_{i1}^x(\theta)\right)}{n_i} & -\frac{p_{i1}^x(\theta)p_{i2}^x(\theta)}{n_i} & \dots & -\frac{p_{i1}^x(\theta)p_{iq-1}^x(\theta)}{n_i} \\ & \frac{p_{i2}^x(\theta)\left(1-p_{i2}^x(\theta)\right)}{n_i} & \dots & -\frac{p_{i2}^x(\theta)p_{iq-1}^x(\theta)}{n_i} \\ & & \dots & \dots \\ & & & \frac{p_{iq-1}^x(\theta)\left(1-p_{iq-1}^x(\theta)\right)}{n_i} \end{bmatrix}$$

for $i = 1 \dots q$ (we assume $n_i = n/q$ is an integer for simplicity).

The estimation strategy now proceeds as explained in Appendix A. The only crucial difference is that we simulate both consumption growth and (residual) income growth using the transition laws

$$\begin{aligned} \ln c_{h,t}^* = \ln c_{h,t-2}^* + \phi\left(\frac{\lambda + r}{1+r}\varepsilon_{h,t} + \frac{(1-\lambda)r}{1+r}\varepsilon_{h,t-1} - \lambda\varepsilon_{h,t-2} + \zeta_{h,t} + \zeta_{h,t-1}\right) \\ + \alpha\left(\upsilon_{h,t} - \upsilon_{h,t-2}\right) \end{aligned}$$

and

$$\begin{aligned} \ln y_{h,t} = \ln y_{h,t-2} + (d_t - d_{t-2}) + \beta(X_{h,t} - X_{h,t-2}) \\ + \zeta_{h,t} + \zeta_{h,t-1} + \varepsilon_{h,t} - \varepsilon_{h,t-2}. \end{aligned}$$

35. We neglect the extra randomness induced by the fact that the class boundaries are preestimated. We also ignore correlation between consumption transition probabilities and income transition probabilities.

References

Alan, Sule, and Martin Browning (2003). "Estimating Euler Equations Using Simulated Residual Estimation." CAM Working Paper 3, University of Copenhagen.

Altonji, Joseph G., and Lewis M. Segal (1996). "Small-Sample Bias in GMM Estimation of Covariance Structures." *Journal of Business and Economic Statistics*, 14, 353–366.

Anderson, Theodore W., and Leo A. Goodman (1957). "Statistical Inference about Markov Chains." *Annals of Mathematical Statistics*, 28, 89–110.

Attanasio, Orazio P. (2000). "Consumption." In *Handbook of Macroeconomics*, edited by John B. Taylor and Michael Woodford. Amsterdam, Elsevier Science.

Attanasio, Orazio P., and Stephen J. Davis (1996). "Relative Wage Movements and the Distribution of Consumption." *Journal of Political Economy*, 104, 1227–1262.

Attanasio, Orazio P., and Tullio Jappelli (2001). "Intertemporal Choice and the Cross-Sectional Variance of Marginal Utility." *Review of Economics and Statistics*, 83, 13–27.

Banerjee, Abhijit V., and Andrew F. Newman (1991). "Risk-Bearing and the Theory of Income Distribution." *Review of Economic Studies*, 58, 211–235.

Bishop, Yvonne M., Stephen E. Fienberg, and Paul W. Holland (1975). *Discrete Multivariate Analysis: Theory and Practice*. Cambridge, MA, MIT Press.

Blundell, Richard W., and Ian Preston (1998). "Consumption Inequality and Income Uncertainty." *Quarterly Journal of Economics*, 113, 603–640.

Blundell, Richard W., Luigi Pistaferri, and Ian Preston (2003). "Partial Insurance, Information, and Consumption Dynamics." CEPR Discussion Paper 3666.

Brandolini, Andrea, and Luigi Cannari (1994). "Methodological Appendix: The Bank of Italy Survey of Household Income and Wealth." In *Saving and the Accumulation of Wealth. Essays on Italian Households and Government Behavior*, edited by A. Ando, L. Guiso, and I. Visco. Cambridge University Press.

Browning, Martin, and Thomas F. Crossley (2001). "The Life-Cycle Model of Consumption and Saving." *Journal of Economic Perspectives*, 15, No. 3, 3–22.

Campbell, John Y., and Gregory N. Mankiw (1989). "Consumption, Income, and Interest Rates: Reinterpreting the Times-Series Evidence." In: *NBER Macroeconomics Annual 1989*, edited by Olivier J. Blanchard and Stanley Fischer, 185–216. MIT Press.

Carroll, Christopher D. (2001). "Precautionary Saving and the Marginal Propensity to Consume out of Permanent Income." NBER Working Paper 8233.

Carroll, Christopher D., and Andrew A. Samwick (1997). "The Nature of Precautionary Wealth." *Journal of Monetary Economics*, 40, 41–72.

Chamberlain, Gary (1984). "Panel Data." In Handbook of Econometrics, vol. 2, edited by Evi Griliches and Michael D. Intriligator. Amsterdam: North-Holland.

Cochrane, John (1991). "A Simple Test of Consumption Insurance." *Journal of Political Economy*, 99, 957–976.

Cutler, David M., and Lawrence F. Katz (1992). "Rising Inequality? Changes in the Distribution of Income and Consumption in the 1980s." *American Economic Review*, 82, 546–551.

Deaton, Angus (1997). *The Analysis of Household Surveys*. Baltimore, The Johns Hopkins University Press.

Deaton, Angus, and Christina Paxson (1994). "Intertemporal Choice and Inequality." *Journal of Political Economy*, 102, 384–394.

Duffle, Darrell, and Kenneth Singleton (1993). "Simulated Moments Estimation of Markov Models of Asset Prices." *Econometrica*, 61, 929–952.

Friedman, Milton (1957). *A Theory of the Consumption Function*. Princeton University Press.

Goffe, William, Gary Ferrier, and John Rogers (1994). "Global Optimization of Statistical Functions with Simulated Annealing." *Journal of Econometrics*, 60, 65–99.

Gourinchas, Pierre-Olivier, and Jonathan Parker (2002). "Consumption Over the Life Cycle." *Econometrica*, 70, 47–91.

Hall, Robert E., and Frederic Mishkin (1982). "The Sensitivity of Consumption to Transitory Income: Evidence from PSID Households." *Econometrica*, 50, 461–481.

Heathcote, Jonathan, Kjetil Storesletten, and Gianluca Violante (2003). "The Welfare Implications of Rising Wage Inequality in the US." Working paper, IIE, Stockholm University.

Jappelli, Tullio, Stephen Pischke, and Nicholas Souleles (1998). "Testing for Liquidity Constraints in Euler Equations with Complementary Data Sources." *The Review of Economics and Statistics*, 80, 251–262.

Krueger, Dirk, and Fabrizio Perri (2003). "On the Welfare Consequences of the Increase in Inequality in the United States." *NBER Macroeconomics Annual 2003*. Cambridge, MA, MIT Press.

Laibson, David (1997). "Golden Eggs and Hyperbolic Discounting." *Quarterly Journal of Economics*, 112, 443–478.

Lee, Bong-Soo, and Beth Ingram (1991). "Simulation Estimation of Time Series Model." *Journal of Econometrics*, 47, 197–205.

Lillard, Lee, and Robert Willis (1978). "Dynamic Aspects of Earning Mobility." *Econometrica*, 46, 985–1012.

Mace, Barbara (1991). "Full Insurance in the Presence of Aggregate Uncertainty." *Journal of Political Economy*, 99, 928–956.

Mas-Colell, Andreu, Michael Whinston, and Jerry Greene (1995). *Microeconomic Theory*. Oxford University Press.

McFadden, Daniel (1989). "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration." *Econometrica*, 57, 995–1026.

Neyman, Jerzy (1949). "Contributions to the Theory of the χ^2 Test." In *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley, University of California Press.

Parker, Jonathan A. (1999). "The Reaction of Household Consumption to Predictable Changes in Social Security Taxes." *American Economic Review*, 89, 959–973.

Phelan, Christopher (1994). "Incentives, Insurance, and the Variability of Consumption and Leisure." *Journal of Economic Dynamics and Control*, 18, 581–599.

Shorrocks, Anthony F. (1978). "The Measurement of Mobility." *Econometrica*, 46, 1013–1024.

Souleles, Nicholas S. (1999). "The Response of Household Consumption to Income Tax Refunds." *American Economic Review*, 89, 947–958.

Townsend, Robert (1994). "Risk and Insurance in Village India." *Econometrica*, 62, 539–591.

Zhang, Qiang, and Masao Ogaki (2001). "Decreasing Risk Aversion and Tests of Risk Sharing." *Econometrica*, 69, 515–526.

This article has been cited by:

1. Katja Kaufmann, Luigi Pistaferri. 2009. Disentangling Insurance and Information in Intertemporal Consumption Choices. *American Economic Review* **99**:2, 387-392. [CrossRef]
2. Tullio Jappelli , Mario Padula , Luigi Pistaferri . 2008. A Direct Test of The Buffer-Stock Model of SavingA Direct Test of The Buffer-Stock Model of Saving. *Journal of the European Economic Association* **6**:6, 1186-1210.